ECF-HANFORD-18-0079
Revision 0

CALCULATION OF CRITICAL MEANS FOR CALENDAR YEAR 2019 RCRA GROUNDWATER MONITORING

Prepared for the U.S. Department of Energy
Assistant Secretary for Environmental Management

Contractor for the U.S. Department of Energy
under Contract DE-AC06-08RL14788

CH2MHILL
Plateau Remediation Company

P.O. Box 1600
Richland, Washington 99352

ECF-HANFORD-18-0079
Revision 0

CALCULATION OF CRITICAL MEANS FOR CALENDAR YEAR 2019 RCRA GROUNDWATER MONITORING

Date Published
May 2019

Prepared for the U.S. Department of Energy
Assistant Secretary for Environmental Management

Contractor for the U.S. Department of Energy
under Contract DE-AC06-08RL14788

CH2MHILL
Plateau Remediation Company
P.O. Box 1600
Richland, Washington 99352

APPROVED
By Julia Raymer at 2:10 pm, May 22, 2019

Release Approval Date

ENVIRONMENTAL CALCULATION COVER PAGE

SECTION 1 - Completed by the Responsible Manager

Project: S&GRP Project Delivery	**RELEASE / ISSUE** DATE: May 22, 2019 HANFORD RELEASE
Date:	
Calculation Title and Description: Calculation of Critical Means for Calendar Year 2019 RCRA Groundwater Monitoring	

Qualifications Summary

Preparer(s):

Name: Erica DiFilippo / SSP&A

Degree, Major, Institution, Year: BA, Geology, Smith College, 2000
MS, Geological Sciences, University of Southern California, 2004
PhD, Hydrology and Water Resources, University of Arizona, 2008

Professional Licenses:

Brief Narrative of Experience: Dr. DiFilippo has extensive experience addressing the fate and transport of organic and inorganic chemicals in the environment. She has conducted numerous studies of groundwater transport, statistical geochemical analysis, groundwater age-dating, dense non-aqueous phase liquid (DNAPL) dissolution, enhanced DNAPL remediation technologies, and the use of innovative technologies for measuring in-situ concentrations of organic contaminants in sediment pore-water. At Hanford, Dr. DiFilippo has provided remediation support in the form of water-level mapping, evaluation of monitored natural attenuation (MNA) and enhanced attenuation (EA), development of methodology to asses the impact of surface water-groundwater interactions on contaminant concentrations and trends, and assessment of statistical methodologies for evaluating RCRA compliance at both interim and final status sites.

ENVIRONMENTAL CALCULATION COVER PAGE (Continued)

Checker(s):

Name: W. Huber / Quant. Decision

Degree, Major, Institution, Year: BA, Mathematics, Haverford College, 1978
BA, Philosophy, Haverford College, 1978
MA, Mathematics, Columbia University, 1979
M. Phil, Mathematics, Columbia University, 1980
PhD, Mathematics, Columbia University, 1985

Professional Licenses:

Brief Narrative of Experience: William A. Huber, PhD, PSTAT, performs general statistical consulting with specialized capabilities in statistical sampling, environmental and spatial statistics, risk analysis, and geographic information systems. He directed the data analysis for the FCC's National Broadband Mapping Plan; was a principal reviewer of the US EPA's guidance document on statistical methods for groundwater monitoring; has testified as a statistical and mathematical modeling expert in complex federal litigation; led research and development of data visualization software; provided statistical support for hundreds of investigations of water, soils, groundwater, wastewater, and air quality; developed custom and commercial statistical software for sampling, spatial data analysis, economic estimation, and database querying; and published research in risk assessment, geostatistics, and quantum mechanics. He has taught mathematics and statistics at St. Joseph’s University, Haverford College, Penn State-Great Valley, Villanova University, and in short courses for professional organizations. For 12 years he taught whitewater kayaking as an ACA-certified Instructor and Instructor Trainer.

Dr. Huber does business as Quantitative Decisions in Rosemont, PA and is also formally associated with Analysis & Inference in Springfield, PA and SS Papadopulos & Associates in Bethesda, MD. At A&I he supports litigation as a consulting and testifying expert. Applications have included assessing rare risks, determining disparate impact in age and gender discrimination cases, and obtaining accurate samples of processes and populations. At SSPA, he provides expertise in statistics and geostatistics to support environmental investigations, remediations, monitoring, and litigation.

Formerly, Dr. Huber owned a software development startup and held senior technical and management positions at a large international engineering consulting firm. Over his career, he has helped develop small businesses and volunteer organizations, serving on their boards of directors in various positions, including as Treasurer of the Philadelphia Canoe Club, Editor of Directions Magazine (a Web publication on GIS), and Director of XI corporation (an environmental services company). He has long been building and nurturing professional communities in GIS and statistics on the Web. Currently he is serving as Treasurer of the Philadelphia Chapter of the American Statistical Association.

ENVIRONMENTAL CALCULATION COVER PAGE (Continued)

Senior Reviewer(s):

Name: R. Murch / INTERA

Degree, Major, Institution, Year: MS, 2002, Civil Engineering, University of South Florida, 2002
BS, 2000, Civil Engineering, University of South Florida, 2000

Professional Licenses: Professional Engineer (Florida)

Brief Narrative of Experience: Renee Murch's professional experience has focused on water resources, hydrology, and civil infrastructure. Her areas of expertise include the development and application of hydraulic, hydrologic, and statistical models to support minimum flow and level development (MFL), restoration of surface water resources, evaluation of saltwater and freshwater interaction, simulation of regional- and local-scale hydrologic conditions as part of water resource planning efforts, and assessment of scouring and erosion processes associated with the construction of bridges and other civil infrastructure. Renee has specialized expertise in the development and application of statistical models including multiple linear regression, artificial neural networks, and Markov Chain Monte Carlo probabilistic simulations. Her work has focused on surface water, groundwater, and integrated modeling applications using modeling applications such as HEC-RAS, MODFLOW, HSPF, ELM, and the Integrated Hydrologic Model. She has evaluated radiological and hydrologic data using methods such as principal component analysis, agglomerative hierarchical cluster analysis, analysis of variance, bivariate correlation, multivariate regression, artificial neural networks, and hypothesis testing. Renee's experience also includes the application of geographic information system tools for data analysis and model input development and hydrologic data collection. She has a variety of field experience related to instrumentation and data collection on both water resources- and geotechnical-related projects, including the installation and maintenance of well transducers, weather stations, evaporation pans, and stream stage, runoff test bed, soil moisture, and tide gauges. Her current work is focused on the development, calibration, and application of surface water, groundwater, and statistical models using applications such as MODFLOW, HEC-RAS, CE-QualW2, ELM, HSPF, IHM, SPLUS, and R to support water supply planning and MFL development.

SECTION 2 - Completed by Preparer

Calculation Number: ECF-HANFORD-18-0079 | Revision Number: 0

Revision History

Revision No.	Description	Date	Affected Pages
0	Initial release		All

SECTION 3 - Completed by the Responsible Manager

Document Control:

Is the document intended to be controlled within the Document Management Control System (DMCS)? ◉ Yes ○ No

Does document contain scientific and technical information intended for public use? ◉ Yes ○ No

Does document contain controlled-use information? ○ Yes ◉ No

ENVIRONMENTAL CALCULATION COVER PAGE (Continued)

SECTION 4 - Document Review and Approval

Preparer(s):			
Erica DiFilippo	Senior Geochemist	Erica DiFilippo	4/8/2019
Print First and Last Name	*Position*	*Signature*	*Date*
Checker(s):			
William Huber	Principal	William A. Huber	4/8/2019 4/8/2019
Print First and Last Name	*Position*	*Signature*	*Date*
Senior Reviewer(s):			
Renee Murch	Sr Water Resources Engr		4/19/19
Print First and Last Name	*Position*	*Signature*	*Date*
Responsible Manager(s):			
William Faught	Manager		5-6-19
Print First and Last Name	*Position*	*Signature*	*Date*

SECTION 5 - Applicable if Calculation is a Risk Assessment or Uses an Environmental Model

Prior to Initiating Modeling:

Required training for modelers completed: Integration Lead:		
N/A		
Print First and Last Name	*Signature*	*Date*
Safety Software Approved: Integration Lead:		
N/A		
Print First and Last Name	*Signature*	*Date*
Calculation Approved: Risk/Modeling Integration Manager:		
N/A		
Print First and Last Name	*Signature*	*Date*

Contents

Appendices

Tables

Terms

CM	critical mean
CR	critical range
CRAN	Comprehensive R Archive Network
CSV	comma-separated values (refers to an ASCII text file format)
CV	coefficient of variation
CY	calendar year
HEIS	Hanford Environmental Information System
KM	Kaplan-Meier
LLWMA	Low-Level Waste Management Area
LOQ	limit of quantitation
RCRA	*Resource Conservation and Recovery Act of 1976*
TOC	total organic carbon
TOX	total organic halides
TXT	text (refers to an ASCII text file format)

This page intentionally left blank.

1 Purpose

This calculation derives the statistical criteria (critical means [CMs]) for indicator parameters used in *Resource Conservation and Recovery Act of 1976* (RCRA) indicator parameter evaluation and detection monitoring at interim status sites and select final status sites for calendar year (CY) 2019. The methodology for calculating the CMs was established in ECF-Hanford-13-0013, *Calculation of Critical Means for Calendar Year 2013 RCRA Groundwater Monitoring*.

2 Background

The Hanford Site is a U.S. Department of Energy facility located in central Washington State along the Columbia River. The site operated as a nuclear production facility starting in the 1940s, and the last reactor was decommissioned in the late 1980s. During operation, many chemicals that can potentially migrate to groundwater were used. Since decommissioning, operations at the Hanford Site have changed to demolition and restoration, including remediation of known groundwater contamination under the *Comprehensive Environmental Response, Compensation, and Liability Act of 1980* and monitoring for new and existing facility impact to groundwater under RCRA. Hanford Site facilities that could create new impacts to groundwater are subject to WAC 173-303, "Dangerous Waste Regulations," which incorporate RCRA regulations for interim status (40 CFR 265, "Interim Status Standards for Owners and Operators of Hazardous Waste Treatment, Storage, and Disposal Facilities") and final status (40 CFR 264, "Standards for Owners and Operators of Hazardous Waste Treatment, Storage, and Disposal Facilities").

As established in WAC 173-303-400, "Interim Status Facility Standards," the facility is required by 40 CFR 265.92(b)(3), "Sampling and Analysis," to monitor pH, specific conductance, total organic carbon (TOC), and total organic halides (TOX) as indicator parameters for groundwater contamination. These parameters are to be measured in four replicate measurements, and a composite measurement should be calculated based on the mean of these four replicates. Monitoring for these parameters is also required by WA7890008967, *Hanford Facility Resource Conservation and Recovery Act Permit, Dangerous Waste Portion, Revision 8C, for the Treatment, Storage, and Disposal of Dangerous Waste* (hereafter referred to as the RCRA Permit), for detection monitoring at certain final-status sites. Samples collected at downgradient wells (compliance wells) are compared to upgradient wells (background wells) to identify potential impacts to groundwater from the facility. If concentrations in the compliance wells are significantly different from CM concentrations in background wells, there is evidence of facility impact to groundwater.

CMs are calculated for the nine RCRA sites listed in Table 1 that are currently under indicator evaluation monitoring. Four RCRA sites previously included in the CM calculations are not included in this calculation:

- The 100-N RCRA sites (1301-N, 1325-N, and 1324-N/NA) were removed from the RCRA Permit in 2018 and RCRA groundwater monitoring requirements for the units are no longer applicable.
- DOE/RL-2013-46, *Groundwater Monitoring Plan for the Liquid Effluent Retention Facility*, was revised in 2017 to change the statistical evaluation method from the Welch's T-Test to the

Double Quantification rule[1] to evaluate waste constituents. The Double Quantification rule, and hence the Liquid Effluent Retention Facility, is not evaluated in this analysis.

This document describes the approach for calculating CMs for statistical comparisons. The calculated CMs will be compared with data collected during CY 2019 to determine if any exceedances (compliance well concentrations higher than the calculated CMs) have occurred. These comparisons will occur in quarterly RCRA monitoring reports and are beyond the scope of this document. CMs are calculated once for each CY unless sampling changes warrant recalculation.

Table 1. RCRA Sites for Indicator Parameter Statistical Comparisons

RCRA Site	Operable Unit
216-A-36B Crib	200-PO-1
216-A-37-1 Crib	200-PO-1
216-B-3 Pond	200-PO-1
216-B-63 Ditch	200-BP-5
216-S-10 Pond and Ditch	200-UP-1
LLWMA-1	200-BP-5
LLWMA-2	200-BP-5
LLWMA-3	200-ZP-1
LLWMA-4	200-ZP-1

Reference: *Resource Conservation and Recovery Act of 1976*, 42 USC 6901, et seq.

LLWMA = Low-Level Waste Management Area

RCRA = *Resource Conservation and Recovery Act of 1976*

2.1 Derivation of and Calculation of Critical Means as Prediction Limits

As established in WAC 173-303-400, the facility is required by 40 CFR 265.92(b)(3) to monitor pH, specific conductance, TOC, and TOX as indicator parameters for groundwater contamination. At each sampling event, these parameters are usually measured in four replicate samples. Replicate measurements are combined into a single numerical result, called a composite measurement (of which each replicate is one aliquot). The composite value is the arithmetic mean of the aliquot values, after first setting the value of any nondetect to its reporting limit. When all the aliquots are nondetects, their composite is considered a nondetect and its reporting limit is set to the smallest reporting limit of the aliquots.

All statistical procedures are carried out separately for each RCRA unit and analyte. The unit-wide false-positive rate is controlled at $\alpha = 1\%$ by means of a Bonferroni adjustment based on the number of compliance decisions anticipated during each monitoring event to which these procedures will apply.

[1] The Double Quantification rule, as defined in EPA 530/R-09-007, 2009, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Unified Guidance*, states that "[a] confirmed exceedance is registered if any well-constituent pair in the '100% non-detect' group exhibits quantified measurements […] in two consecutive sample and resample events."

A set of reference (background) data is created for each analyte at each RCRA unit. In most cases, these data are obtained at designated upgradient wells during a specified (recent) period. The reference data, which consist of replicate samples obtained at each well during each sampling event, are composited into one value per well per sampling event. This background dataset is used to compute CMs (or for pH, a critical range [CR]). These are prediction limits (or intervals, respectively) for data that will be obtained at downgradient (compliance) wells during any single monitoring event. Composite measurements from those downgradient wells will be compared to CMs and CRs. Any downgradient composite result that exceeds its CMs (or lies outside its CRs, in the case of pH) will be considered statistically significant evidence that the groundwater it represents differs from conditions represented by the reference dataset. Such a determination is called a positive result.

CMs and CRs are derived from the test statistic for the Average Replicate t-test and calculated using the following formulas:

Critical mean
$$CM = m + t_{df,k,\alpha}\, s \sqrt{1 + \frac{1}{n}} \qquad \textbf{(Eq. 2.1)}$$

Critical range
$$CR = m \pm t_{df,k,\alpha/2}\, s \sqrt{1 + \frac{1}{n}} \qquad \textbf{(Eq. 2.2)}$$

where:

m = estimates the background mean

s = estimates the background standard deviation

n = the count of data used to compute the mean m

α = the minimum unit-wide false-positive rate for any single (future) monitoring event; set at 1%

k = the number of comparisons (counting all analytes) that have the potential to create a positive result within a RCRA unit during any single monitoring event

$t_{df,k,\alpha}$ = the upper 100% – α/k quantile of Student’s t distribution with df degrees of freedom

$t_{df,k,\alpha/2}$ = the upper 100% – α/(2k) quantile of Student’s t distribution with df degrees of freedom

df = the degrees of freedom (equal to n – 1 for the interwell tests)

The Student *t* quantiles have been generically referred to as “t_{crit}”. The quantities α/k and $\alpha/(2k)$ have been referred to as the “adjusted” α or $\alpha_{adjusted}$. These terms are used in the tabulated results appearing in Chapter 7. The Student *t*-test was used in accordance with 40 CFR 265.93, “Preparation, Evaluation, and Response.”

Typically, *k* will equal the number of upgradient and downgradient wells multiplied by the number of analytical parameters being tested (always equal to 4).

According to ECF-Hanford-13-0013, CMs (and CRs) are recalculated annually or if the number of comparisons changes. Annual recalculation, using a sliding period for selecting the background data, is intended to account for changing background conditions as provided in Section 5.3.5.3 of the *Statistical Guidance for Ecology Site Managers* (Washington State Depart of Ecology publication number 92-54). Changes in the number of comparisons are usually the result of changes in monitoring well networks (i.e., wells are added or deleted). If changes occur in a monitoring well network, CMs and CRs for that facility are recalculated for subsequent sampling events using the new well network.

2.2 Counting Comparisons for the Bonferroni Adjustment

For comparability with previous calculations, as set forth in ECF-Hanford-13-0013, the number of comparisons for this analysis includes both upgradient and downgradient wells. Under 40 CFR 265.93(b), the comparison must consider individually each of the wells in the monitoring system. As such, at each RCRA unit, the number of analytes is multiplied by the total number of wells, including background wells, to accommodate planned testing of future results at upgradient wells.

3 Methodology

CMs were calculated using the interwell approach following procedures outlined in EPA 530/R-09-007, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Unified Guidance*, to set limits based on data from an upgradient well that is assumed to be unimpacted by activities at the monitored site.

3.1 Background and Compliance Wells

At each RCRA site, background wells are identified as wells that were hydraulically upgradient of the facility. Compliance wells are identified as wells that were hydraulically downgradient of the facility. A list of wells and well locations based on RCRA sites is presented in Appendix A.

3.2 Data Acquisition

Groundwater chemistry data are downloaded from the Hanford Environmental Information System (HEIS), which is maintained by CH2M HILL Plateau Remediation Company, and exported into a Microsoft® Access® database (HEIS_Chem1_12192018.accdb and HEIS_Chem2_12192018.accdb). Data for this analysis were downloaded from the HEIS database on December 19, 2018. The HEIS database contains one table (HEIS2_ADM_PNLGW_STD_RESULT_MV_1 and HEIS2_ADM_PNLGW_STD_RESULT_MV_2, respectively), which contains information on groundwater chemistry samples, including lab and review data qualifiers, sample medium, sample collection purpose, analytical method, and reporting limits. Fields extracted from the HEIS database for use in the CM calculations are presented in Table 2.

The RCRA parameter data (pH, specific conductance, TOC, and TOX) from the HEIS database are exported into two text (TXT) files named qryChemHeis1.txt and qryChemHeis2.txt.

Table 2. HEIS Database Fields

Field Extracted	Definition
WELL_NAME	Location ID
SAMP_DATE_TIME	Sampling Date
STD_CON_LONG_NAME	Analyte Name
STD_VALUE_RPTD	Reported Concentration

® Microsoft and Access are registered trademarks of Microsoft Corporation in the United States and other countries.

Table 2. HEIS Database Fields

Field Extracted	Definition
STD_ANAL_UNITS_RPTD	Units for Concentration Measurement
LAB_QUALIFIER	Lab Data Qualifier
REVIEW_QUALIFIER	Review Data Qualifier [a]
COLLECTION_PURPOSE	Primary Reason for Sample Collection [b]
VALIDATION_QUALIFIER	Validation Qualifier

a. Review Data Qualifier codes are as follows:

F = The result is undergoing further review.
G = Record has been reviewed and determined to be correct, or the record has been modified to make it correct.
H = Laboratory holding time exceeded before the sample was analyzed.
P = Potential problem. Collection/analysis circumstances make value questionable.
Q = Associated quality control sample is out of limits.
R = Do not use. Further review indicates the result is not valid. These points are not included in outputs or downloads.
Y = Result suspect. Review – insufficient evidence to show result valid or invalid.
Z = Miscellaneous circumstances exist. Additional information may be found in the RESULT_COMMENT field for this record.

b. Sample collection purpose codes are as follows:

C = Characterization
IH = Industrial Hygiene
IP = In Process
R = Routine
S = Special Studies
T = Transportation
TS = Time Series
VER = Verification
VP = Vertical Profile
WM = Waste Management

HEIS = Hanford Environmental Information System

3.3 Review Qualifiers

Data are removed from the dataset prior to calculation of CMs based on their review qualifiers (Table 3). This step was conducted for each individual sample prior to compiling composite samples.

Table 3. Review Qualifiers for Data Removal

Review Qualifier	Definition
Y	Result is suspect. Review had insufficient evidence to show result valid or invalid.
R	Do not use. Further review indicates the result is not valid.

3.4 Collection Purpose

Data for the calculation of CMs are limited to data that were collected for routine purposes or verification (i.e., COLLECTION_PURPOSE = "R" or "VER"). This step was conducted for each individual sample prior to compiling composite samples.

3.5 Outliers

Based on visual inspection of the data, obvious outliers were removed prior to analysis (Table 4). The removal of this outlier is consistent with previous years' calculations (ECF-Hanford-18-0004, *Calculation of Critical Means for Calendar Year 2018 RCRA Groundwater Monitoring*). Additional outlier testing was performed on composite samples (see Section 3.12) to identify any other potential outliers.

Table 4. Outliers Removed from Analysis

Well	Contaminant of Concern	Sample Date	Basis
699-26-38	Total organic carbon	10/17/2016	An order of magnitude higher than all other measured concentrations at this location

3.6 Calculation of Composite Results

At each RCRA site and for each well and analyte, composite results are calculated as the average of the replicates (its aliquot measurements). When none of the aliquots in the composite (also referred to as a "composite sample" to emphasize the composite nature of the measurements) are classified as a nondetect, the following statistics are calculated to characterize the background conditions and to support the CM/CR calculations:

Estimated background mean for composite sample

$$m_c = \frac{1}{n_c}\sum_{i=1}^{n_c} x_i \quad \textbf{(Eq. 3.1)}$$

Estimated standard deviation for composite sample

$$s_c = \sqrt{\frac{1}{n_c - 1}\sum_{i=1}^{n_c}(x_i - m_c)^2} \quad \textbf{(Eq. 3.2)}$$

where:

m_c = the composite sample mean

n_c = the number of aliquots in the composite sample

x_i = the aliquot measurement

s_c = the composite sample standard deviation

When one or more values in the dataset are nondetects, the procedures described in Section 3.9 (handling nondetects) are used to compute m_c and s_c.

3.7 Identifying Background Data

CMs for each RCRA site are calculated from measurements at wells identified as being upgradient of the RCRA site (i.e., background wells). Classification of well location (upgradient versus downgradient) is presented in Appendix A. Changes to groundwater remedies may likely change groundwater flow at the 216-B-63 Ditch and Low-Level Waste Management Area 2 (LLWMA) RCRA sites. Table 5 presents the current background well networks identified in the groundwater monitoring plans and proposed background wells based on evaluation of groundwater flow changes from ongoing groundwater remedies. CMs are calculated for both the current well network and proposed changes. For RCRA areas not presented in Table 5, no changes were made to the well networks.

Table 5. RCRA Areas with Multiple Calculated CMs/CRs

RCRA Area	Network	Background Wells
216-B-63 Ditch	Current	299-E33-33 299-E34-12 299-E34-8
	Proposed	299-E27-19 299-E33-33 299-E34-12 299-E34-8
LLWMA-2	Current	299-E34-2
	Proposed	299-E27-10 299-E34-2 299-E34-9

CM = critical mean
CR = critical range
LLWMA = low level waste management area
RCRA = *Resource Conservation and Recovery Act of 1976*

3.8 Determination of Date Range

EPA 530/R-09-007 (pg. 5-3) recommends at least "8 to 10 independent background observations be collected before running most statistical tests." To ensure this, the number of composite measurements for each analyte at each RCRA site is counted over a minimum 3-year period covering 2016 and 2018 and adjusted, as necessary, to provide an adequate background dataset. Whenever this count is less than eight for any analyte at a RCRA site, the date range is expanded backwards in time until all analytes have at least eight samples. This common date range is used to estimate background conditions for the CR/CM calculations at that site. If the number of detected values within the date range for TOC or TOX was less than two, earlier data was added for that indicator parameter so that there was a minimum of two nondetects in the dataset.

Sites with adjustments to the date range beyond the three-year period or eight sample minimum are:

- 216-B-3 Pond: Date range expanded to 10/16/2015 to reduce the percent non-detects to $\leq$80% for TOX.
- 216-B-63 Ditch: Date range expanded to 4/3/2015 to reduce the percent non-detects to $\leq$80% for TOX.
- 216-S-10 Pond and Ditch: Sample size is restricted to four due to documented spatial variability in specific conductance (ECF-200UP1-12-0055, *Analysis of Spatial Variability of Specific Conductance in Groundwater at the 216-S-10 Pond & Ditch*). To maintain a sample size of four for pH, specific conductance, and TOX, a start date of analysis for CM calculations was set at May 2, 2017. The date range for TOC included data starting from May 7, 2014, so that the TOC dataset contained at least two detected values.
- LLWMA-2: Date range expanded to 4/3/2015 to reduce the percent non-detects to $\leq$80% for TOC.

3.9 Handling Nondetects

Processing of nondetects proceeds through two stages. In the first stage, raw data are combined into composite measurements. Each composite is flagged according to the detectability of the data contributing to its value. These composites are grouped into reference datasets for each analyte at each RCRA unit. At the second stage, estimates of the means (m) and standard deviations (s) required for the CM/CR calculations are computed, depending on what proportion of the reference dataset consists of nondetects.

3.9.1 Stage 1: Computing Composite Values

All decisions in this monitoring program are based on composite measurements. That is, when four measurements are available at one monitoring event for an analyte at a well, these measurements are called replicates and are combined into a single value called a composite measurement for that event-analyte-well combination. The replicate measurements become the aliquots of the resulting composite measurement.

At the outset, beginning with the data selected as described in Sections 3.2 through 3.8, the following summary statistics are computed for each composite, c. The individual aliquot values are denoted x_i with indexes ranging from 1 through n_c. When replicate i is not detected, x_i equals the reporting limit associated with that sample.

Number of nondetects	$n_{c,0}$	**(Eq. 3.3)**
Proportion of nondetects	$\frac{n_{c,0}}{n_c}$	**(Eq. 3.4)**
Minimum detection limit[2]	min (x_i | result i not detected)	**(Eq. 3.5)**
Maximum detected value[3]	max(x_i)	**(Eq. 3.6)**

[2] Defined as detection limit in ECF-Hanford-13-0013, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Unified Guidance* .

[3] When all aliquots of a composite sample are nondetects, the maximum detected value is equal to the maximum reporting limit.

Mean of all values $$m_c = \frac{1}{n_c}\sum_{i=1}^{n_c} x_i$$ **(Eq. 3.7)**

Coefficient of variation (CV) (calculated when all aliquots of a composite sample are detected values) $$\frac{s_{c,d}}{m_{c,d}}$$ **(Eq. 3.8)**

where:

$s_{c,d}$ = the standard deviation of the composite sample when all aliquots are detected (i.e., no nondetects)

$m_{c,d}$ = the mean of the composite sample when all aliquots are detected (i.e., no nondetects)

The composite measurement is taken to be the mean of all values (m_c), and the composite measurement standard deviation is s_c. Multiple reporting limits were not present for composite sample aliquots (i.e., all aliquots of any composite sample had the same reporting limit).

The reporting limit is substituted for aliquots with measurements below the reporting limit. When all of the aliquots are nondetects, the sample is treated as a nondetect in the subsequent analysis and its reporting limit is set to the smallest of the aliquot reporting limits. If the proportion of the nondetects is less than 1, the composite measurement is treated as a detected value in the subsequent analysis.

3.9.2 Stage 2: Treatment of Nondetects for Computing Statistics of Reference Datasets

The treatment of nondetects for datasets with less than 10 percent nondetects is discussed in Section 3.9.2.1, and datasets with greater than 10 percent nondetects are discussed in Section 3.9.2.2.

3.9.2.1 Datasets with Less Than 10 Percent Nondetects

EPA 530/R-09-007 (pg. 15-5) indicates that t-tests and prediction limits (and therefore the CM/CR calculations that are derived from them) are not significantly affected by substitution of half the reporting limit when the proportion of nondetects is no more than 10 percent to 15 percent of the total sample. Therefore, this substitution method is employed here for datasets where up to 10 percent of the composite samples are identified as nondetects. Equations 3.1 and 3.2 are then applied as if all values were actually quantified.

3.9.2.2 Datasets with Greater than 10 Percent Nondetects

The Kaplan-Meier (KM) estimator is used to account for nondetects in datasets where more than 10 percent of the composite measurements are identified as nondetects. The KM method is a nonparametric approach that uses a ranked ordering method to estimate the proportion of concentrations below each reporting limit. EPA 530/R-09-007 provides a detailed description of the KM method (pg. 15-7). The calculations are carried out using the EnvStats package in R (see Chapter 5 of this document). These calculations produce estimates of the background mean and standard deviation, replacing those in Equations 3.1 and 3.2.

3.10 Test for Normality

The background datasets are tested at the $\alpha = 0.05$ level using the Shapiro-Wilk test to evaluate the null hypothesis that the data are normally distributed. EPA 530/R-09-007 (pg. 10-13) provides a detailed description of the Shapiro-Wilk test. It is conducted using the “stats” package in R (see Chapter 5 of this document). CMs and CRs calculated from datasets that show evidence of departing from normal distributions are flagged in the tabulated results presented in Chapter 7. The Shapiro-Wilk test was also used to evaluate the null hypothesis that the data are log-normally distributed.

3.11 Test for Outliers

The background datasets are tested using the Grubbs test at the $\alpha = 0.001$ level to identify outliers, consistent with ECF-Hanford-13-0013. This test is useful for identifying exactly one outlier in a dataset (Grubbs, 1969, "Procedures for Detecting Outlying Observations in Samples") using composite measurements. The Grubbs test is conducted using the "outliers" package in R (see Chapter 5 of this document). Outliers were removed at the discretion of project scientists. No additional outliers to those listed in Table 4 were removed for the CY 2019 CM calculations based on the Grubbs test results.

3.12 Intrawell Method

Following the procedures set forth in ECF-Hanford-18-0004, an intrawell test was used for the LLWMA-2 and LLWMA-3 sites. The LLWMA-3 background and compliance monitoring wells may be affected by injection of treated effluent from the 200 West Pump and Treat, which is injected downgradient of the background well used for establishing background at LLWMA-3.

For upgradient-to-downgradient well comparisons, a crucial assumption is that downgradient well changes in groundwater quality are only caused by onsite releases of regulated constituents. The influence of the injection of treated effluent result in spatial variability potentially leading to large numbers of false-positive detections at the compliance wells based on CMs calculated by the interwell method. Intrawell testing is identified in EPA 530/R-09-007 as a method to eliminate the problem of natural spatial variability.

Water level mapping and particle-tracking simulations described in a draft engineering evaluation report for low-level burial ground Trench 94 (in process) indicate the LLWMA-2 upgradient monitoring well is not upgradient, but is cross-gradient. Upgradient wells are not feasible due to the basalt outcropping above the water table upgradient of the site. For LLWMA-2, intrawell testing background may therefore provide more accurate baseline for use in statistical comparisons[4].

Interim status monitoring regulations applicable to LLWMA-2 and LLWMA-3 do not provide for intrawell testing; however, CMs are calculated using the intrawell method and provided for comparison in Chapter 7.

Intrawell limit methods use previous results collected in a monitoring well to set a limit that determines whether future samples in the same well represent an adverse change in groundwater quality. A statistically significant ("positive") result is interpreted as evidence of a real difference between the previous results and the future measurement.

The initial steps for obtaining and subsetting data for the intrawell method are the same as described in Sections 3.2 through 3.8. The method for handling nondetects and assessing normality in the intrawell method are the same as described in Sections 3.9 and 3.10. Calculations specific to the intrawell method are described in the following subsections.

The wells evaluated using the intrawell test are presented in Table 6.

[4] In the calculations, well 299-E34-2 is identified as an upgradient well.

Table 6. RCRA Sites and Wells Evaluated Using the Intrawell Test

RCRA Area	Wells
LLWMA-2	299-E27-8 299-E27-9 299-E27-10 299-E27-11 299-E27-17 299-E34-2 299-E34-9 299-E34-10 299-E34-12
LLWMA-3	299-W9-2 299-W10-29 299-W10-30 299-W10-31

LLWMA = low level waste management area

RCRA = *Resource Conservation and Recovery Act of 1976*

3.12.1 Test for Homogeneity of Variance

By default, the variances in the logarithms of the composite indicator parameter results are assumed to be the same in each well. This assumption is tested using Levene's test (adjusted for nondetects). EPA 530/R-09-007 (pg. 11-4) provides a detailed description of Levene's test. Levene's test is conducted using the "lawstat" package in R (see Chapter 5 of this document) at an α = 1% level. CMs and CRs calculated from datasets that show evidence of departing from the assumption of having the same variance are flagged in the tabulated results presented in Chapter 7. Individual well means together with a single pooled standard deviation are used for computing the CMs.

3.12.2 Calculation of Test Statistics and Critical Means

As established in ECF-HANFORD-13-0013, the CM should be calculated as in Equations 2.1 and 2.2, but based on a pooled standard deviation (s_{ij}) estimated from all the compliance wells in the RCRA unit. The wells, which will be indexed by j = 1, 2, 3, contribute three reference datasets (x_{ij}) where j (which indexes the individual composite results) ranges from 1 through n_j (the amount of reference data available at each well). The total count of reference data is written n. For specific conductance, TOC, and TOX, let $y_{ij} = \ln(x_{ij})$ and for pH, let $y_{ij} = x_{ij}$. Let n_w be the number of wells in the pooled dataset (in this case, 3).

Individual well log means $$m_j = \frac{1}{n_j}\sum_{i=1}^{n_j} y_{ij}$$ **(Eq. 3.10)**

Individual well log variance $$s_j^2 = \frac{1}{n_j - 1}\sum_{i=1}^{n_j}\left(y_{ij} - m_j\right)^2$$ **(Eq. 3.11)**

Pooled log variance $$s^2 = \frac{1}{\sum_{j=1}^{n_w}(n_j - 1)}\sum_{j=1}^{n_w}(n_j - 1)\, s_j^2 = \frac{1}{n-3}\sum_{j=1}^{n-3}(n_j - 1)s_j^2$$ **(Eq. 3.12)**

Pooled well standard deviation	$s = \sqrt{s^2}$	**(Eq. 3.13)**
Pooled degrees of freedom	$df_{pooled} = n - n_w$	**(Eq. 3.14)**
Individual well critical means	$CM_j = \exp\left(m_j + t_{df_{pooled},k,\alpha} s \sqrt{1+\frac{1}{n_j}} \right)$	**(Eq. 3.15)**
Individual well critical range	$CM_j = m_j \pm t_{df_{pooled},k,\alpha} s \sqrt{1+\frac{1}{n_j}}$	**(Eq. 3.16)**

The value of k is set, as in all other cases, to the total number of analytes (four).

Each intrawell dataset was evaluated to determine if the pooled standard deviation approach was appropriate. Datasets were evaluated for temporal trends and spread using trend analysis, violin plots of the raw data, and violin plots of median absolute deviation and range in standard deviation. A summary of the pooled standard deviation evaluation is presented in Table 7.

Table 7. Summary of Pooled Standard Deviation Evaluation

RCRA Area	Analyte	Use Pooled Standard Deviation?	Basis
LLWMA-2	pH	Yes	
LLWMA-2	Specific conductance	No	Several wells display temporal trends
LLWMA-2	Total organic carbon	No	Majority of wells have > 50% nondetects
LLWMA-2	Total organic halides	No	Spatial variability among wells
LLWMA-3	pH	Yes	
LLWMA-3	Specific conductance	Yes	
LLWMA-3	Total organic carbon	No	Spatial variability among wells
LLWMA-3	Total organic halides	No	Several wells display temporal trends

LLWMA = low level waste management area

RCRA = *Resource Conservation and Recovery Act of 1976*

4 Assumptions and Inputs

The following assumptions are made in the CMs analysis:

1. Reference data from upgradient wells are assumed to represent natural groundwater or background conditions.
2. Unless otherwise noted, it is assumed that the mean and variance of each analytical parameter exhibit no spatial variation within each region represented by the upgradient wells at any RCRA unit.
3. Time series of results are assumed stationary, which implies they have no trend. While it is recognized that groundwater quality may naturally change with time, accounting for trends is deliberately avoided by choosing a relative short period of records and periodically recalculating CMs.
4. The Student's t-test, from which the CM and CR limits are derived, assumes the sampling distribution of the mean follows a normal distribution and the sampling distribution of the variance independently is proportional to a chi-squared distribution with the assumed degrees of freedom. A large coefficient of variation (CV) can be an indication of non-normality of the mean (Equation 3.8). As a diagnostic adjunct to the Shapiro-Wilk test of normality, the CV is calculated each time a CM is calculated.
5. It is assumed that no seasonal fluctuations occur in the data.
6. Typically, datasets with large percentages of nondetects (e.g., greater than 50%) will result in CMs close to the detection limits but less than quantitation limits (which are approximately three times the detection limits). When the CM is below the limit of quantitation (LOQ), the LOQ is also used as a comparison value for detecting facility effects. LOQ calculations and comparisons are made elsewhere.

Table 8 summarizes the number of data used in the calculation of CMs/CRs, including the number of data removed by each data processing step.

Table 8. Summary of Data used for Calculation of CMs/CRs

	Number of Samples	Percent of Total
Number of Total Aliquots from HEIS Database		
Total Number of Data from HEIS	**144,903**	
pH	57,599	40%
Specific conductance	58,147	40%
TOC	16,044	11%
TOX	13,113	9.0%
Number of Total Aliquots for Upgradient Wells		
Total Number of Data for Upgradient Wells	**10,483**	
pH	2,605	25%
Specific conductance	2,743	26%

Table 8. Summary of Data used for Calculation of CMs/CRs

	Number of Samples	Percent of Total
TOC	2,604	25%
TOX	2,531	24%
Number of Aliquots Removed Based on Review Qualifiers	**115**	**1.1%**
pH	4	0.15%
Specific conductance	1	0.04%
TOC	62	2.4%
TOX	48	1.9%
Number of Aliquots Removed Based on Collection Purpose	**8**	**0.08%**
pH	0	0.0%
Specific conductance	0	0.0%
TOC	8	0.31%
TOX	0	0.0%
Number of Aliquots	**10,360**	**99%**
Number of Nondetects	**2,486**	**24%**
pH	0	0.0%
Specific conductance	0	0.0%
TOC	1,080	41%
TOX	1,406	56%
Number of Composite Measurements for Upgradient Wells		
Number of Event-Analyte-Well Combinations	**2,159**	
pH	581	22%
Specific conductance	616	22%
TOC	483	19%
TOX	479	19%

CM = critical mean
CR = critical range
HEIS = Hanford Environmental Information System
TOC = total organic carbon
TOX = total organic halides

5 Software Applications

CMs were calculated using the public domain computing platform R (version 3.4.3 published 11/30/2017). R provides data manipulation, calculation, and graphical display capabilities to support data analysis (Venables et al., 2016, *An Introduction to R Notes on R: A Programming Environment for Data Analysis and Graphics*). It is freely available to the public and can be compiled and run on a variety of platforms (UNIX, Windows, and Mac OS). The base installation of R contains statistical and plotting functions and many more are available for download through the Comprehensive R Archive Network (CRAN). Several R packages are used for calculating CMs. The packages were downloaded from the CRAN and are listed in Table 9.

Table 9. R Packages Used for Critical Means Calculations

[R] Package	Package Description	Version
data.table	Enhanced data processing	1.10.4-3
EnvStats	Package for Environmental Statistics, Including US EPA Guidance	2.3.0
lawstat	Tools for biostatistics, public policy, and law	3.2
outliers	Tests for outliers	0.14
plyr	Tools for splitting, applying, and combining data	1.8.4
reshape2	Restructure and aggregate data	1.4.3

EPA = U.S. Environmental Protection Agency

6 Calculation

The R code used to process the data and calculate CMs for each site is included in Appendix B. The text files (qryChemHeis1.txt and qryChemHeis2.txt) contain the data downloaded from the HEIS database (see Section 3.2), and the comma-separated values (csv) file (ClosureSamples_01302017.csv) contains the list of closure samples to be excluded from the analysis. In addition, a csv file containing information on well location for each RCRA site (RCRA_Indicator_Monitoring_Wells.csv) and a csv file containing the start dates for analysis (Start_Dates.csv) are used in the R code.

Output information is written to the files:

1. HANFORD-CRITICAL_MEANS-INTERWELL_2019-(date of analysis).csv – Contains CMs calculated using the Interwell Method
2. HANFORD-CRITICAL_MEANS-INTRAWELL_2019-(date of analysis).csv – Contains CMs for LLWMA-2 and LLWMA-3 calculated using the Intrawell Method
3. DataSummary_ALLHEISSamples_(date of analysis).txt – Contains all of the data removed based on review qualifiers
4. DataSummary_AllUpgradient_(date of analysis).txt – Contains a summary of the total number of samples for upgradient wells
5. Totals_Removed_Qualifier_(date of analysis).txt – Contains a summary of the total number of samples removed based on the review qualifier

6. Totals_Removed_Collection_(date of analysis).txt – Contains a summary of the total number of samples removed based on the collection purpose
7. Totals_Aliquots-(date of analysis).txt – Contains a summary of the total number of aliquots in the final dataset
8. Totals_NDs-(date of analysis).txt – Contains a summary of the total number of nondetects in the final dataset
9. Totals_Composite-(date of analysis).txt – Contains a summary of the total number of composite samples in the final dataset

This information has been formatted to produce the tables presented in the results section of this report as well as in the annual groundwater monitoring report.

7 Results

Results are listed in Tables 10 through 22.

Table 10. Critical Means for the 216-A-36B Crib for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	2			
Date Range for Data	1/4/2016 – 7/13/2018			
Number of Composite Samples	19	16	16	19
% Nondetects	0%	44%	25%	0%
Mean	656 µS/cm	336 µg/L	4.79 µg/L	7.84 su
Median	669 µS/cm	331 µg/L	3.74 µg/L	7.87 su
Standard Deviation	68.8 µS/cm	224 µg/L	3.01 µg/L	0.258 su
Coefficient of Variation	0.105	0.668	0.63	0.0328
Maximum	773 µS/cm	1,110 µg/L	15.7 µg/L	8.16 su
Number of Comparisons	24	24	24	24
alpha per comp ($\alpha_{adjusted}$)	4.17E-04	4.17E-04	4.17E-04	2.08E-04
$t_{crit,i}$	4.00	4.16	4.16	4.32
Critical Mean	938 µS/cm	1,300 [a,b] µg/L	17.7 [a,b] µg/L	8.98 [a,b] su
Lower Critical Range				6.70 [a,b] su

a. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

b. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

CY = calendar year

su = standard unit

Table 11. Critical Means for the 216-A-37-1 Crib for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	2			
Date Range for Data	1/7/2016 – 10/8/2018			
Number of Composite Samples	21	21	21	21
% Nondetects	0%	24%	62%	0%
Mean	470 µS/cm	366 µg/L	4.44 µg/L	8.30 su
Median	502 µS/cm	361 µg/L	<3.64 [a] µg/L	8.24 su
Standard Deviation	60.6 µS/cm	84.5 µg/L	2.15 µg/L	0.165 su
Coefficient of Variation	0.129	0.231	0.485	0.0199
Maximum	545 µS/cm	545 µg/L	10.4 µg/L	8.75 su
Number of Comparisons	24	24	24	24
alpha per comp ($\alpha_{adjusted}$)	4.17E-04	4.17E-04	4.17E-04	2.08E-04
$t_{crit,i}$	3.93	3.93	3.93	4.22
Critical Mean	714 [b,c] µS/cm	705 µg/L	13.1 [b,c] µg/L	9.01 su
Lower Critical Range				7.58 su

a. Determination of the median includes a high percentage of non-detects at the reporting limit.

b. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

c. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

CY = calendar year

su = standard unit

Table 12. Critical Means for the 216-B-3 Pond for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	3			
Date Range for Data	10/16/2015 – 10/2/2018			
Number of Composite Samples	20	18	18	20
% Nondetects	0%	22%	78%	0%
Mean	303 µS/cm	411 µg/L	3.75 µg/L	8.03 su
Median	284 µS/cm	411 µg/L	<3.33 [a] µg/L	8.03 su
Standard Deviation	37.8 µS/cm	133 µg/L	1.04 µg/L	0.153 su
Coefficient of Variation	0.125	0.324	0.279	0.019
Maximum	363 µS/cm	671 µg/L	7.80 µg/L	8.25 su
Number of Comparisons	20	20	20	20
alpha per comp ($\alpha_{adjusted}$)	5.00E-04	5.00E-04	5.00E-04	2.50E-04
$t_{crit,i}$	3.88	3.97	3.97	4.19
Critical Mean	453 [b,c] µS/cm	953 µg/L	8.00 [b,c,d] µg/L	8.69 [b,c] su
Lower Critical Range				7.38 [b,c] su

a. Determination of the median includes a high percentage of non-detects at the reporting limit.

b. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

c. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

d. Percent nondetects in dataset used to calculate critical mean is greater than 50%, therefore the statistical comparison values are the larger of either the CM or the LOQ.

CM = critical mean

CY = calendar year

LOQ = limit of quantitation

su = standard unit

Table 13. Critical Means for the 216-B-63 Ditch for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	3			
Date Range for Data	4/3/2015 – 11/5/2018			
Number of Composite Samples	25	25	24	25
% Nondetects	0%	44%	75%	0%
Mean	628 µS/cm	304 µg/L	3.52 µg/L	8.12 su
Median	648 µS/cm	330 µg/L	<3.33 [a] µg/L	8.15 su
Standard Deviation	71.4 µS/cm	99.1 µg/L	0.491 µg/L	0.101 su
Coefficient of Variation	0.114	0.326	0.14	0.0125
Maximum	707 µS/cm	720 µg/L	7.70 µg/L	8.35 su
Number of Comparisons	24	24	24	24
alpha per comp ($\alpha_{adjusted}$)	4.17E-04	4.17E-04	4.17E-04	2.08E-04
$t_{crit,i}$	3.82	3.82	3.84	4.09
Critical Mean	907 [b,c] µS/cm	690 [b,c] µg/L	5.45 [b,c,d] µg/L	8.55 su
Lower Critical Range				7.70 su

a. Determination of the median includes a high percentage of non-detects at the reporting limit.

b. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

c. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

d. Percent nondetects in dataset used to calculate critical mean is greater than 50%, therefore the statistical comparison values are the larger of either the CM or the LOQ.

CM = critical mean

CY = calendar year

LOQ = limit of quantitation

su = standard unit

Table 14. Critical Means for the 216-B-63 Ditch - New for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	4			
Date Range for Data	4/3/2015 – 11/5/2018			
Number of Composite Samples	35	34	32	34
% Nondetects	0%	47%	66%	0%
Mean	640 µS/cm	298 µg/L	3.56 µg/L	8.14 su
Median	660 µS/cm	330 µg/L	<3.39 [a] µg/L	8.16 su
Standard Deviation	66.5 µS/cm	94.9 µg/L	1.23 µg/L	0.143 su
Coefficient of Variation	0.104	0.318	0.345	0.0175
Maximum	736 µS/cm	720 µg/L	8.18 µg/L	8.53 su
Number of Comparisons	24	24	24	24
alpha per comp ($\alpha_{adjusted}$)	4.17E-04	4.17E-04	4.17E-04	2.08E-04
$t_{crit,i}$	3.67	3.68	3.70	3.92
Critical Mean	887 [b,c] µS/cm	652 [b,c] µg/L	8.17 [b,c,d] µg/L	8.71 su
Lower Critical Range				7.58 su

a. Determination of the median includes a high percentage of non-detects at the reporting limit.

b. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

c. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

d. Percent nondetects in dataset used to calculate critical mean is greater than 50%, therefore the statistical comparison values are the larger of either the CM or the LOQ.

CM = critical mean

CY = calendar year

LOQ = limit of quantitation

su = standard unit

Table 15. Critical Means for the 216-S-10 Pond and Ditch for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	1			
Date Range for Data	5/2/2017 – 11/14/2018	5/7/2014 – 11/14/2018	5/2/2017 – 11/14/2018	
Number of Composite Samples	4	10	4	4
% Nondetects	0%	80%	25%	0%
Mean	304 µS/cm	229 µg/L	4.33 µg/L	7.51 su
Median	305 µS/cm	<350 [a] µg/L	3.72 µg/L	7.51 su
Standard Deviation	6.14 µS/cm	83.0 µg/L	2.10 µg/L	0.0483 su
Coefficient of Variation	0.0202	0.362	0.485	0.00643
Maximum	309 µS/cm	720 µg/L	7.78 µg/L	7.57 su
Number of Comparisons	20	20	20	20
alpha per comp ($\alpha_{adjusted}$)	5.00E-04	5.00E-04	5.00E-04	2.50E-04
$t_{crit,i}$	12.9	4.78	12.9	16.3
Critical Mean	392 µS/cm	645 [c] µg/L	34.6 [b] µg/L	8.39 su
Lower Critical Range				6.63 su

a. Determination of the median includes a high percentage of non-detects at the reporting limit.

b. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

c. Percent nondetects in dataset used to calculate critical mean is greater than 50%, therefore the statistical comparison values are the larger of either the CM or the LOQ.

CM = critical mean

CY = calendar year

LOQ = limit of quantitation

su = standard unit

Table 16. Critical Means for Low-Level Waste Management Area 1 for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	2			
Date Range for Data	1/15/2016 – 7/20/2018			
Number of Composite Samples	11	11	11	11
% Nondetects	0%	45%	45%	0%
Mean	433 µS/cm	456 µg/L	3.65 µg/L	7.94 su
Median	435 µS/cm	400 µg/L	3.33 µg/L	7.91 su
Standard Deviation	20.8 µS/cm	160 µg/L	2.17 µg/L	0.118 su
Coefficient of Variation	0.048	0.351	0.595	0.0149
Maximum	471 µS/cm	763 µg/L	8.28 µg/L	8.10 su
Number of Comparisons	28	28	28	28
alpha per comp ($\alpha_{adjusted}$)	3.57E-04	3.57E-04	3.57E-04	1.79E-04
$t_{crit,i}$	4.81	4.81	4.81	5.28
Critical Mean	538 µS/cm	1,260 [a,b] µg/L	14.6 [a,b] µg/L	8.59 su
Lower Critical Range				7.29 su

a. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

b. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

CY = calendar year

su = standard unit

Table 17. Critical Means for Low-Level Waste Management Area 2 for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	1			
Date Range for Data	4/3/2015 – 11/2/2018			
Number of Composite Samples	9	9	8	9
% Nondetects	0%	56%	0%	0%
Mean	1,140 µS/cm	559 µg/L	10.9 µg/L	7.80 su
Median	1,120 µS/cm	<720 [a] µg/L	10.2 µg/L	7.82 su
Standard Deviation	83.4 µS/cm	240 µg/L	4.09 µg/L	0.0502 su
Coefficient of Variation	0.0731	0.429	0.376	0.00644
Maximum	1,340 µS/cm	1,020 µg/L	16.8 µg/L	7.84 su
Number of Comparisons	36	36	36	36
alpha per comp ($\alpha_{adjusted}$)	2.78E-04	2.78E-04	2.78E-04	1.39E-04 su
$t_{crit,i}$	5.53	5.53	5.98	6.14
Critical Mean	1,630 [b] µS/cm	1,960 µg/L	36.8 µg/L	8.12 su
Lower Critical Range				7.47 su

a. Determination of the median includes a high percentage of non-detects at the reporting limit.

b. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

CY = calendar year

su = standard unit

Table 18. Critical Means for Low-Level Waste Management Area 2 – New for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	3			
Date Range for Data	4/3/2015 – 11/2/2018			
Number of Composite Samples	30	26	25	30
% Nondetects	0%	77%	12%	0%
Mean	933 µS/cm	437 µg/L	7.16 µg/L	7.92 su
Median	1,090 µS/cm	<649 [a] µg/L	6.70 µg/L	7.93 su
Standard Deviation	285 µS/cm	182 µg/L	4.01 µg/L	0.112 su
Coefficient of Variation	0.305	0.416	0.56	0.0141
Maximum	1,420 µS/cm	1,020 µg/L	16.8 µg/L	8.15 su
Number of Comparisons	36	36	36	36
alpha per comp ($\alpha_{adjusted}$)	2.78E-04	2.78E-04	2.78E-04	1.39E-04
$t_{crit,i}$	3.88	3.96	3.98	4.13
Critical Mean	2,060 [b,c] µS/cm	1,170 [b,c] µg/L	23.4 µg/L	8.39 su
Lower Critical Range				7.45 su

a. Determination of the median includes a high percentage of non-detects at the reporting limit.

b. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

c. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

CY = calendar year

su = standard unit

Table 19. Intrawell Critical Means for Low Level Waste Management Area 2 for CY 2019 Comparisons

Statistic	299-E34-2	299-E27-10	299-E27-11	299-E27-17	299-E27-8	299-E27-9	299-E34-10	299-E34-12	299-E34-9
Specific Conductance Intrawell Critical Means									
Mean	575 µS/cm	1,150 µS/cm	496 µS/cm	505 µS/cm	482 µS/cm	1,060 µS/cm	679 µS/cm	553 µS/cm	1,010 µS/cm
Standard Deviation	11.6 µS/cm	82.3 µS/cm	14.0 µS/cm	7.79 µS/cm	25.8 µS/cm	68.8 µS/cm	33.9 µS/cm	82.1 µS/cm	259 µS/cm
Number of Composite Samples	8	8	10	10	8	8	12	8	13
% Non-Detects	0%	0%	0%	0%	0%	0%	0%	0%	0%
Maximum	597 µS/cm	1,340 µS/cm	518 µS/cm	515 µS/cm	537 µS/cm	1,190 µS/cm	734 µS/cm	699 µS/cm	1,420 µS/cm
alpha	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$t_{crit,i}$	3.00	3.00	2.82	2.82	3.00	3.00	2.72	3.00	2.68
Critical Mean	612 µS/cm	1,410 µS/cm	537 µS/cm	528 µS/cm	564 µS/cm	1,280 µS/cm	775 µS/cm	815 µS/cm	1,730 µS/cm
Total Organic Carbon Intrawell Critical Means									
Mean	355 µg/L	566 µg/L	240 µg/L	364 µg/L	256 µg/L	743 µg/L	301 µg/L	307 µg/L	422 µg/L
Standard Deviation	62.3 µg/L	230 µg/L	98.7 µg/L	133 µg/L	98.1 µg/L	164 µg/L	65.6 µg/L	69.7 µg/L	121 µg/L
Number of Composite Samples	10	11	10	10	12	11	10	10	12
% Non-Detects	80%	55%	60%	60%	83%	9%	60%	50%	75%
Maximum	720 µg/L	1,020 µg/L	720 µg/L	763 µg/L	720 µg/L	961 µg/L	720 µg/L	424 µg/L	738 µg/L
alpha	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$t_{crit,i}$	2.82	2.76	2.82	2.82	2.72	2.76	2.82	2.82	2.72
Critical Mean	539 * µg/L	1,230 µg/L	533 * µg/L	759 * µg/L	534 * µg/L	1,220 µg/L	495 * µg/L	513 µg/L	764 * µg/L
Total Organic Halides Intrawell Critical Means									
Mean	7.02 µg/L	11.1 µg/L	5.93 µg/L	6.24 µg/L	8.21 µg/L	5.93 µg/L	4.28 µg/L	3.36 µg/L	4.10 µg/L
Standard Deviation	2.60 µg/L	4.34 µg/L	2.65 µg/L	3.40 µg/L	3.55 µg/L	3.88 µg/L	1.75 µg/L	0.0444 µg/L	1.85 µg/L

Table 19. Intrawell Critical Means for Low Level Waste Management Area 2 for CY 2019 Comparisons

Statistic	299-E34-2	299-E27-10	299-E27-11	299-E27-17	299-E27-8	299-E27-9	299-E34-10	299-E34-12	299-E34-9
Number of Composite Samples	8	7	7	8	7	7	8	8	9
% Non-Detects	13%	0%	14%	25%	0%	14%	38%	75%	22%
Maximum	11.2 µg/L	16.8 µg/L	10.2 µg/L	12.9 µg/L	15.3 µg/L	14.6 µg/L	7.70 µg/L	7.70 µg/L	7.63 µg/L
alpha	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$t_{crit,i}$	3.00	3.14	3.14	3.00	3.14	3.14	3.00	3.00	2.90
Critical Mean	15.3 µg/L	25.7 µg/L	14.9 µg/L	17.1 µg/L	20.1 µg/L	19.0 µg/L	9.84 µg/L	3.50 * µg/L	9.74 µg/L
pH Measurement Intrawell Critical Means									
Mean	8.05 su	7.79 su	8.15 su	8.03 su	8.10 su	8.05 su	8.00 su	8.08 su	7.93 su
Standard Deviation	0.0576 su	0.0529 su	0.178 su	0.153 su	0.0786 su	0.10 su	0.0838 su	0.124 su	0.0609 su
Number of Composite Samples	8	8	9	10	8	8	12	8	13
% Non-detects	0%	0%	0%	0%	0%	0%	0%	0%	0%
Pooled Standard Deviation (log-scale)	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106
Pooled Sample Size	84	84	84	84	84	84	84	84	84
Maximum	8.15 su	7.84 su	8.40 su	8.19 su	8.25 su	8.20 su	8.11 su	8.24 su	8.02 su
alpha	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
df_{pooled}	75	75	75	75	75	75	75	75	75
$t_{crit,i}$	2.64	2.64	2.64	2.64	2.64	2.64	2.64	2.64	2.64
Upper Critical Range	8.33 su	8.08 su	8.43 su	8.32 su	8.38 su	8.33 su	8.28 su	8.36 su	8.21 su
Lower Critical Range	7.77 su	7.51 su	7.87 su	7.75 su	7.81 su	7.77 su	7.71 su	7.79 su	7.64 su

* Percent nondetects in dataset used to calculate critical mean is greater than 50%, therefore the statistical comparison values are the larger of either the CM or the LOQ.

CM = critical mean

Table 19. Intrawell Critical Means for Low Level Waste Management Area 2 for CY 2019 Comparisons

Statistic	299-E34-2	299-E27-10	299-E27-11	299-E27-17	299-E27-8	299-E27-9	299-E34-10	299-E34-12	299-E34-9

CY = calendar year

LOQ = limit of quantitation

su = standard unit

Table 20. Critical Means for Low-Level Waste Management Area 3 for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	1			
Date Range for Data	3/2/2015 – 9/10/2018			
Number of Composite Samples	8	8	8	8
% Nondetects	0%	25%	38%	0%
Mean	393 µS/cm	367 µg/L	4.36 µg/L	8.06 su
Median	389 µS/cm	351 µg/L	3.62 µg/L	8.07 su
Standard Deviation	11.9 µS/cm	54.5 µg/L	1.25 µg/L	0.0986 su
Coefficient of Variation	0.0302	0.148	0.286	0.0122
Maximum	409 µS/cm	507 µg/L	6.51 µg/L	8.18 su
Number of Comparisons	16	16	16	16
alpha per comp ($\alpha_{adjusted}$)	6.25E-04	6.25E-04	6.25E-04	3.13E-04
$t_{crit,i}$	5.20	5.20	5.20	5.86
Critical Mean	459 µS/cm	668 [a,b] µg/L	11.2 [a,b] µg/L	8.67 su
Lower Critical Range				7.45 su

a. Dataset not normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

b. Dataset not log normally distributed based on Shapiro-Wilk Test ($p < 0.05$).

CY = calendar year

su = standard unit

Table 21. Intrawell Critical Means for Low/ Level Waste Management Area 3 for CY 2019 Comparisons

Statistic	299-W9-2	299-W10-29	299-W10-30	299-W10-31
Specific Conductance Intrawell Critical Means				
Mean	393 µS/cm	405 µS/cm	397 µS/cm	487 µS/cm
Standard Deviation	11.9 µS/cm	8.74 µS/cm	20.1 µS/cm	14.1 µS/cm
Number of Composite Samples	8	8	9	9
% Non-Detects	0%	0%	0%	0%
Pooled Standard Deviation (log scale)	0.0349	0.0349	0.0349	0.0349
Pooled Sample Size	34	34	34	34
Maximum	409 µS/cm	416 µS/cm	430 µS/cm	505 µS/cm
alpha	0.01	0.01	0.01	0.01
df_{pooled}	30	30	30	30
$t_{crit,i}$	2.46	2.46	2.46	2.46
Critical Mean	429 µS/cm	441 µS/cm	433 µS/cm	531 µS/cm
Total Organic Carbon Intrawell Critical Means				
Mean	367 µg/L	307 µg/L	309 µg/L	281 µg/L
Standard Deviation	54.5 µg/L	202 µg/L	75.5 µg/L	181 µg/L
Number of Composite Samples	8	8	8	8
% Non-Detects	25%	63%	50%	75%
Maximum	507 µg/L	720 µg/L	463 µg/L	720 µg/L
alpha	0.01	0.01	0.01	0.01
$t_{crit,i}$	3.00	3.00	3.00	3.00
Critical Mean	540 µg/L	949 * µg/L	549 µg/L	855 * µg/L
Total Organic Halides Intrawell Critical Means				
Mean	3.74 µg/L	7.51 µg/L	11.2 µg/L	13.0 µg/L
Standard Deviation	2.00 µg/L	2.74 µg/L	3.48 µg/L	5.83 µg/L
Number of Composite Samples	8	8	9	9
% Non-Detects	9%	9%	9%	9%
Maximum	6.51 µg/L	12.9 µg/L	16.4 µg/L	22.0 µg/L

Table 21. Intrawell Critical Means for Low/ Level Waste Management Area 3 for CY 2019 Comparisons

Statistic	299-W9-2	299-W10-29	299-W10-30	299-W10-31
alpha	0.01	0.01	0.01	0.01
$t_{crit,i}$	3.00	3.00	2.90	2.90
Critical Mean	10.1 µg/L	16.2 µg/L	21.8 µg/L	30.8 µg/L
pH Measurement Intrawell Critical Means				
Mean	8.06 su	7.97 su	7.92 su	7.86 su
Standard Deviation	0.0986 su	0.0851 su	0.0848 su	0.17 su
Number of Composite Samples	8	8	9	9
% Non-Detects	0%	0%	0%	0%
Pooled Standard Deviation (log scale)	0.117	0.117	0.117	0.117
Pooled Sample Size	34	34	34	34
Maximum	8.18 su	8.09 su	8.04 su	8.06 su
alpha	0.01	0.01	0.01	0.01
df_{pooled}	30	30	30	30
$t_{crit,i}$	2.75	2.75	2.75	2.75
Upper Critical Range	8.38 su	8.30 su	8.24 su	8.19 su
Lower Critical Range	7.73 su	7.65 su	7.59 su	7.54 su

* Percent nondetects in dataset used to calculate critical mean is greater than 50%, therefore the statistical comparison values are the larger of either the CM or the LOQ.

CM = critical mean

CY = calendar year

LOQ = limit of quantitation

su = standard unit

Table 22. Critical Means for Low-Level Waste Management Area 4 for CY 2019 Comparisons

Statistic	Specific Conductance	Total Organic Carbon	Total Organic Halides	pH Measurement
Number of Upgradient Wells	1			
Date Range for Data	1/17/2014 – 6/15/2018			
Number of Composite Samples	10	8	9	10
% Nondetects	0%	63%	0%	0%
Mean	598 µS/cm	329 µg/L	11.1 µg/L	8.09 su
Median	598 µS/cm	<385 * µg/L	8.95 µg/L	8.10 su
Standard Deviation	20.5 µS/cm	54.1 µg/L	6.02 µg/L	0.186 su
Coefficient of Variation	0.0344	0.164	0.541	0.023
Maximum	626 µS/cm	720 µg/L	22.1 µg/L	8.37 su
Number of Comparisons	24	24	24	24
alpha per comp ($\alpha_{adjusted}$)	4.17E-04	4.17E-04	4.17E-04	2.08E-04
$t_{crit,i}$	4.91	5.58	5.19	5.43
Critical Mean	704 µS/cm	649 µg/L	44.1 µg/L	9.15 su
Lower Critical Range				7.03 su

* Determination of the median includes a high percentage of non-detects at the reporting limit.

CY = calendar year

su = standard unit

8 References

40 CFR 264, "Standards for Owners and Operators of Hazardous Waste Treatment, Storage, and Disposal Facilities," *Code of Federal Regulations*. Available at: http://www.gpo.gov/fdsys/pkg/CFR-2010-title40-vol25/xml/CFR-2010-title40-vol25-part264.xml.

40 CFR 265, "Interim Status Standards for Owners and Operators of Hazardous Waste Treatment, Storage, and Disposal Facilities," *Code of Federal Regulations*. Available at: http://www.gpo.gov/fdsys/pkg/CFR-2010-title40-vol25/xml/CFR-2010-title40-vol25-part265.xml.

265.92, "Sampling and Analysis."

265.93, "Preparation, Evaluation, and Response."

Comprehensive Environmental Response, Compensation, and Liability Act of 1980, 42 USC 9601, et seq., Pub. L. 107-377, December 31, 2002. Available at: https://www.csu.edu/cerc/researchreports/documents/CERCLASummary1980.pdf.

DOE/RL-2013-46, 2017, *Groundwater Monitoring Plan for the Liquid Effluent Retention Facility*, Rev. 1, U.S. Department of Energy, Richland Operations Office, Richland, Washington. Available at: https://pdw.hanford.gov/arpir/pdf.cfm?accession=0068832H.

ECF-200UP1-12-0055, 2012, *Analysis of Spatial Variability of Specific Conductance in Groundwater at the 216-S-10 Pond & Ditch*, Rev. 1, CH2M HILL Plateau Remediation Company, Richland, Washington. Available at: https://pdw.hanford.gov/arpir/pdf.cfm?accession=0065957H.

ECF-Hanford-13-0013, 2014, *Calculation of Critical Means for Calendar Year 2013 RCRA Groundwater Monitoring*, Rev. 1, CH2M HILL Plateau Remediation Company, Richland, Washington. Available at: https://pdw.hanford.gov/arpir/pdf.cfm?accession=0080432H.

ECF-HANFORD-18-0004, 2018, *Calculation of Critical means for Calendar Year 2018 RCRA Groundwater Monitoring*, Rev. 0, CH2M HILL Plateau Remediation Company, Richland, Washington. Available at: https://pdw.hanford.gov/arpir/pdf.cfm?accession=0064440H.

Ecology Publication 92-54, 1992, *Statistical Guidance for Ecology Site Managers*, Toxics Cleanup Program, Washington State Department of Ecology, Olympia, Washington. Available at: https://fortress.wa.gov/ecy/publications/documents/9254.pdf.

EPA 530/R-09-007, 2009, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Unified Guidance*, Office of Resource Conservation and Recovery, U.S. Environmental Protection Agency, Washington, D.C. Available at: https://www.itrcweb.org/gsmc-1/Content/Resources/Unified_Guidance_2009.pdf.

Grubbs, Frank E., 1969, "Procedures for Detecting Outlying Observations in Samples," *Technometrics* 11(1):1-21. Available at: http://web.ipac.caltech.edu/staff/fmasci/home/astro_refs/OutlierProc_1969.pdf.

Resource Conservation and Recovery Act of 1976, 42 USC 6901, et seq. Available at: https://elr.info/sites/default/files/docs/statutes/full/rcra.pdf.

Venables, W.N., D.M. Smith, and the R Core Team, 2016, *An Introduction to R Notes on R: A Programming Environment for Data Analysis and Graphics*, Version 3.3.1, 95 pp.

WA7890008967, 2007, *Hanford Facility Resource Conservation and Recovery Act Permit, Dangerous Waste Portion, Revision 8C, for the Treatment, Storage, and Disposal of Dangerous Waste*, as amended, Washington State Department of Ecology, Richland, Washington. Available at: https://fortress.wa.gov/ecy/nwp/permitting/hdwp/rev/8c/.

WAC 173-303, "Dangerous Waste Regulations," *Washington Administrative Code*, Olympia, Washington. Available at: http://apps.leg.wa.gov/WAC/default.aspx?cite=173-303.

303-400, "Interim Status Facility Standards."

This page intentionally left blank.

Appendix A

List of Upgradient and Downgradient Wells by Site

This page intentionally left blank.

A1 List of Upgradient and Downgradient Wells by Site

Table A-1 presents the list of RCRA Site upgradient and downgradient wells used for the critical mean calculations.

Table A-1. List of Wells by Site

Site [a]	Well Name [b]	Location [c]	Use [d]
216-A-36B Crib	299-E17-1	UP	TRUE
216-A-36B Crib	299-E17-14	DOWN	TRUE
216-A-36B Crib	299-E17-15	DOWN	TRUE
216-A-36B Crib	299-E17-16	DOWN	TRUE
216-A-36B Crib	299-E17-18	DOWN	TRUE
216-A-36B Crib	299-E17-19	UP	TRUE
216-A-37-1 Crib	299-E25-17	DOWN	TRUE
216-A-37-1 Crib	299-E25-19	DOWN	TRUE
216-A-37-1 Crib	299-E25-20	DOWN	TRUE
216-A-37-1 Crib	299-E25-35	UP	TRUE
216-A-37-1 Crib	299-E25-47	UP	TRUE
216-A-37-1 Crib	299-E25-95	DOWN	TRUE
216-B-3 Pond	699-42-42B	DOWN	TRUE
216-B-3 Pond	699-43-43B	DOWN	FALSE
216-B-3 Pond	699-43-44	DOWN	FALSE
216-B-3 Pond	699-43-45	DOWN	TRUE
216-B-3 Pond	699-44-39B	UP	TRUE
216-B-3 Pond	699-44-43C	UP	TRUE
216-B-3 Pond	699-45-42	UP	TRUE
216-B-63 Ditch	299-E27-16	DOWN	TRUE
216-B-63 Ditch	299-E27-18	DOWN	TRUE
216-B-63 Ditch	299-E27-19	DOWN	TRUE
216-B-63 Ditch	299-E33-33	UP	TRUE
216-B-63 Ditch	299-E34-12	UP	TRUE
216-B-63 Ditch	299-E34-8	UP	TRUE
216-B-63 Ditch-New	299-E27-16	DOWN	TRUE
216-B-63 Ditch-New	299-E27-18	DOWN	TRUE
216-B-63 Ditch-New	299-E27-19	UP	TRUE
216-B-63 Ditch-New	299-E33-33	UP	TRUE

Table A-1. List of Wells by Site

Site [a]	Well Name [b]	Location [c]	Use [d]
216-B-63 Ditch-New	299-E34-12	UP	TRUE
216-B-63 Ditch-New	299-E34-8	UP	TRUE
216-S-10 Pond and Ditch	299-W26-13	DOWN	TRUE
216-S-10 Pond and Ditch	299-W26-14	DOWN	TRUE
216-S-10 Pond and Ditch	299-W27-2	DOWN	FALSE
216-S-10 Pond and Ditch	699-32-76	DOWN	TRUE
216-S-10 Pond and Ditch	699-33-75	DOWN	TRUE
216-S-10 Pond and Ditch	699-33-76	UP	TRUE
LLWMA-1	299-E28-26	DOWN	TRUE
LLWMA-1	299-E28-27	DOWN	TRUE
LLWMA-1	299-E28-28	DOWN	TRUE
LLWMA-1	299-E28-33	DOWN	FALSE
LLWMA-1	299-E32-3	UP	TRUE
LLWMA-1	299-E33-266	UP	TRUE
LLWMA-2	299-E27-10	UP	TRUE
LLWMA-2	299-E27-11	DOWN	TRUE
LLWMA-2	299-E27-17	DOWN	TRUE
LLWMA-2	299-E27-8	DOWN	TRUE
LLWMA-2	299-E27-9	DOWN	TRUE
LLWMA-2	299-E34-10	DOWN	TRUE
LLWMA-2	299-E34-12	DOWN	TRUE
LLWMA-2	299-E34-2	DOWN	TRUE
LLWMA-2	299-E34-9	DOWN	TRUE
LLWMA-2-New	299-E27-10	UP	TRUE
LLWMA-2-New	299-E27-11	DOWN	TRUE
LLWMA-2-New	299-E27-17	DOWN	TRUE
LLWMA-2-New	299-E27-8	DOWN	TRUE
LLWMA-2-New	299-E27-9	DOWN	TRUE
LLWMA-2-New	299-E34-10	DOWN	TRUE
LLWMA-2-New	299-E34-12	DOWN	TRUE
LLWMA-2-New	299-E34-2	UP	TRUE
LLWMA-2-New	299-E34-9	UP	TRUE

Table A-1. List of Wells by Site

Site [a]	Well Name [b]	Location [c]	Use [d]
LLWMA-3	299-W10-29	DOWN	TRUE
LLWMA-3	299-W10-30	DOWN	TRUE
LLWMA-3	299-W10-31	DOWN	TRUE
LLWMA-3	299-W9-2	UP	TRUE
LLWMA-4	299-W15-152	DOWN	TRUE
LLWMA-4	299-W15-17	DOWN	FALSE
LLWMA-4	299-W15-224	DOWN	TRUE
LLWMA-4	299-W15-30	DOWN	TRUE
LLWMA-4	299-W15-83	DOWN	TRUE
LLWMA-4	299-W15-94	DOWN	TRUE
LLWMA-4	299-W17-1	UP	FALSE
LLWMA-4	299-W18-21	UP	TRUE
LLWMA-4	299-W18-22	UP	FALSE
LLWMA-4	299-W18-40	DOWN	FALSE

a. Site is the RCRA site.

b. Well name is the name of the well.

c. Location is the specified location (upgradient or downgradient) of the well from the RCRA site.

d. “TRUE” indicates data from the well are used in statistical calculations; “FALSE” indicates they are not.

LLWMA = low level waste management area

RCRA = *Resource Conservation and Recovery Act of 1976*

This page intentionally left blank.

Appendix B

Critical Means Analysis R Code

This page intentionally left blank.

B1 Critical Means Analysis R Code

Appendix B includes the R code used to process the data and calculate critical means for each site.

```
#------------------------------------------------------------------------------#
#--This script imports data for the RCRA Critical Means Calculations - 2019
#--By: E. DiFilippo (SSP&A)
#--R Version: 3.1.3
#--Package data.table version: 1.9.4
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Set working directory--#
DIR <- 'Z:/1639_HANFORD_FY2019/66007-23 RCRA Regulatory Support GW Data Evaluation/RCRA Critical Means'
setwd(DIR)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Load Libraries--#
library(data.table)
library(XLConnect)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import User-Defined Functions--#
source('R_Functions/readCHEMHEIS.R')
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Set Site-Wide Error Rate--#
ALPHA <- 0.01
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import Chemistry Data from HEIS--#
DATA1 <- readCHEMHEIS('Data/Data Pull - 12-19-2018/qryChemHEIS1.txt')
DATA2 <- readCHEMHEIS('Data/Data Pull - 12-19-2018/qryChemHEIS2.txt')
DATA <- rbind(DATA1,DATA2)
DATA$EVENT <- as.Date(DATA$EVENT)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import RCRA Site and Wells List--#
RCRA <- loadWorkbook('Data/Received from Art Lee - 12-27-2018/RCRA_Wells_for_2019_CM_Calcs - Revised 3-7-2019.xlsx')
RCRASITES <- readWorksheet(RCRA,sheet = 'Well_Network',header = TRUE)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Update Well List per e-mail from Art Lee on 12/10/2018--#
RCRASITES$Use <- ifelse(RCRASITES$Well.Name == '699-45-42',TRUE,RCRASITES$Use.for.Statistical.Evaluation)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Subset Date After 2004--#
DATA <- DATA[EVENT >= as.Date(ISOdate(2004,01,01))]
DATA <- DATA[EVENT < as.Date(ISOdate(2019,01,01))]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Subset by Analyte--#
AN <- c('pH Measurement','Specific Conductance','Total organic halides','Total organic carbon')
DATA <- subset(DATA,DATA$ANALYTE %in% AN)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import RCRA Site Closure Samples--#
CLOS <- fread('Data/ClosureSamples_01302017.csv')
CLOS$EVENT <- as.Date(as.POSIXct(CLOS$Date_Collected,format='%m/%d/%Y',tz='UTC'))
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Remove Closure Samples--#
ROW <- which(DATA$NAME %in% CLOS$Well & month(DATA$EVENT) == 12 & year(DATA$EVENT) == 2016)
DATA <- DATA[!ROW,]
```

```
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Remove TOC Outlier--#
ROW <- which(DATA$NAME == '699-26-38' & DATA$ANALYTE == 'Total organic carbon' &
           DATA$EVENT == as.Date(ISOdate(2016,10,17)))
DATA <- DATA[!ROW,]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Drop wells identified but not used--#
RCRASITES <- subset(RCRASITES, Use == TRUE )
RCRASITES <- subset(RCRASITES,!RCRA.Site == 'LERF')
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Identify nonDetects; TRUE indicates nondetect--#
DATA$NDS <-  ifelse(DATA$LABQ %like% 'U',TRUE,FALSE)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Identify Data to Remove Based on Review Qualifiers--#
DATA$REVIEW_RM <- ifelse(DATA$REVIEWQ %like% 'Y' | DATA$REVIEWQ %like% 'R',TRUE,FALSE)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--use only data from 'routine' or verification collection purpose--#
DATA$COLLECT_RM <- ifelse(DATA$COLLECTION_PURPOSE %in% c("R",'VER'),FALSE,TRUE)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Create Intrawell Datasets--#
INTRASITES <- subset(RCRASITES,RCRASITES$RCRA.Site %in% c('LLWMA-2','LLWMA-3'))
INTRASITES_SC <- subset(RCRASITES,RCRASITES$RCRA.Site %in% c('116-N-3 (1325-N)','120-N-1 and 120-N-2 (1324-N/NA)'))

INTRA <- subset(DATA,DATA$NAME %in% INTRASITES$Well.Name)
INTRA_SC <- subset(DATA,DATA$NAME %in% INTRASITES_SC$Well.Name & DATA$ANALYTE=='Specific Conductance')
INTRA <- rbind(INTRA,INTRA_SC)
INTRA <- list(INTRASITES = rbind(INTRASITES,INTRASITES_SC),
          DATA = INTRA)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Create Interwell Datasets--#
ROWS <- which(DATA$NAME %in% INTRASITES_SC$Well.Name & DATA$ANALYTE=='Specific Conductance')
INTER <- DATA[!ROWS,]
INTER <- list(INTERSITES = subset(RCRASITES,RCRASITES$Well.Name %in% INTER$NAME),
          DATA = INTER)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Save Initial Datasets--#
save(INTRA,file='Data/Runs - 3-7-2019/Intrawell.RData')
save(INTER,file='Data/Runs - 3-7-2019/Interwell.RData')
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--This script calculates the Critical Means for Hanford RCRA Areas for CY-2019
#--Interwell Method
#--Script written by: E. DiFilippo (SSP&A): modified from ECF-HANFORD-13-0013, REV. 0
#--R Version: 3.1.3
#--Package data.table version: 1.10.4.2
#--Package EnvStats version: 2.1.0
#--Package outliers version: 0.14
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Set working directory--#
DIR <- 'Z:/1639_HANFORD_FY2019/66007-23 RCRA Regulatory Support GW Data Evaluation/RCRA Critical Means'
setwd(DIR)
```

```
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Load Libraries--#
library(data.table)
library(EnvStats)
library(outliers)
library(XLConnect)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import Interwell Dataset--#
load('Data/Runs - 3-7-2019/Interwell.RData')
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Set Site-Wide Error Rate--#
ALPHA <- 0.01
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import Staring Dates--#
RCRA <- loadWorkbook('Data/Received from Art Lee - 12-27-2018/RCRA_Wells_for_2019_CM_Calcs - Revised 3-7-2019.xlsx')
SDATES <- readWorksheet(RCRA,sheet = 'Start_Dates',header = TRUE)
SDATES$START_DATE <- as.Date(as.POSIXct(SDATES$START_DATE,format='%m/%d/%Y',tz='UTC'))
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Subset for Upgradient Wells--#
UP <- as.data.table(subset(INTER$INTERSITES,INTER$INTERSITES$Loc == 'UP'))
DATA <- subset(INTER$DATA,INTER$DATA$NAME %in% UP$Well.Name)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Remove Data Based on Review Qualifiers and Collection Purpose--#
DATA <- DATA[REVIEW_RM == FALSE]
DATA <- DATA[COLLECT_RM == FALSE]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Merge Sample Data and Well Info Data--#
setkey(DATA,NAME)
setkey(UP,Well.Name)
DATA <- UP[DATA,allow.cartesian=TRUE]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Check units consistancy--#
UNITCHECK <- aggregate(DATA$UNIT, by=list(DATA$ANALYTE),
              function(x){length(unique(x))})
if(max(UNITCHECK$x) > 1){
  stop("Different units - check")
}
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Calculate the number of comparison made at each site--#
NCOMP <- data.table(as.data.frame(table(INTER$INTERSITES$RCRA.Site)))
setnames(NCOMP,colnames(NCOMP),c("Site","NCOMP"))
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Create lists of sites and analyte names--#
SITES <- sort(unique(DATA$RCRA.Site))
ANALYTES <- unique(DATA$ANALYTE)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Calculating Summary Statistics for each Well by Sample Date
DATA_SUM <- data.table(DATA)[,list(N_NDS=sum(NDS),
```

```
                              N_SAMPLE=length(NDS),
                              PROPND=sum(NDS)/length(NDS),
                              MEAN_REPORTED=mean(VAL),
                              DETLIMIT=min(VAL),
                              MEAN_DETECTS=mean(VAL[!NDS]),
                              SD_DETECTS=sd(VAL[!NDS]),
                              CV_DETECTS=sd(VAL[!NDS])/mean(VAL[!NDS])),
                         by=c("RCRA.Site","Loc","Well.Name","EVENT","ANALYTE")]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Identify nonDetects; TRUE indicates nondetect for Summary Dataset--#
DATA_SUM$NDS <- ifelse(DATA_SUM$PROPND == 1,TRUE,FALSE)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Calculate Summary Statistics for Each Site including NDs--#
DATA_SUM_COUNTS <- data.table(DATA_SUM)[,list(ALL_NDS=sum(PROPND==1),
                                         N_SAMPLES=length(PROPND),
                                         PROPNDALL=round(100*sum(PROPND==1)/length(PROPND))),
                                    by=c("ANALYTE","RCRA.Site")]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Creating data frame to place results for exporting--#
OUT <- NULL
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Loop through all Sites--#
for(i in 1:length(SITES)) {

  #------------------------------------------------------------------------------#
  #--Subsetting dataset by Site--#
  SUB <- DATA_SUM[RCRA.Site == SITES[i]]
  #------------------------------------------------------------------------------#

  #------------------------------------------------------------------------------#
  #--Number of comparisons--#
  NCOMP_SUB <- subset(NCOMP,Site == SITES[i])$NCOMP * length(ANALYTES)
  #------------------------------------------------------------------------------#

  #------------------------------------------------------------------------------#
  #--Analyte Loop--#
  for(j in 1:length(ANALYTES)){

    #------------------------------------------------------------------------------#
    #--Subset Data by Analyte--#
    ANALYTE_SUB <- SUB[ANALYTE == ANALYTES[j]]
    #------------------------------------------------------------------------------#

    #------------------------------------------------------------------------------#
    #--Subset data by start date--#
    STARTDATE <- subset(SDATES,SDATES$RCRA.Site == SITES[i] & SDATES$ANALYTE == ANALYTES[j])
    ANALYTE_SUB <- subset(ANALYTE_SUB,ANALYTE_SUB$EVENT >= STARTDATE$START_DATE)
    #------------------------------------------------------------------------------#

    #------------------------------------------------------------------------------#
    #--Skip Analytes with No Data--#
    if(nrow(ANALYTE_SUB) == 0){
      next
    }
    #------------------------------------------------------------------------------#

    #------------------------------------------------------------------------------#
    #--Handling Non-detects--#
    if(mean(ANALYTE_SUB$NDS) <= 0.1){

      ND_METHOD <- "HALF.RL"
```

```
    ANALYTE_SUB$ADJVAL <- ifelse(ANALYTE_SUB$NDS ==
TRUE,ANALYTE_SUB$MEAN_REPORTED/2,ANALYTE_SUB$MEAN_REPORTED)
    MN <- mean(ANALYTE_SUB$ADJVAL)
    SD <- sd(ANALYTE_SUB$ADJVAL)

   } else {

    ND_METHOD <- "Kaplan-Meier"

    KM <- enparCensored(ANALYTE_SUB$MEAN_REPORTED,
               ANALYTE_SUB$NDS)
    MN <- as.numeric(KM$parameters[1])
    SD <- as.numeric(KM$parameters[2])

   }
   #--------------------------------------------------------------------------#

   #--------------------------------------------------------------------------#
   #--Determining Critical Values for Student t-Test--#
   if(ANALYTES[j] == 'pH Measurement'){
    ALPHA_CRIT <- (ALPHA/NCOMP_SUB)/2
   }      else {
    ALPHA_CRIT <- ALPHA/NCOMP_SUB
   }

   N <- nrow(ANALYTE_SUB)
   TCRIT <- qt(ALPHA_CRIT,df=N-1,lower.tail=FALSE,log=FALSE)
   CMUP <- MN + TCRIT *SD * sqrt(1 + 1/N)
   CMDW <- MN - TCRIT *SD * sqrt(1 + 1/N)
   #--------------------------------------------------------------------------#

   #--------------------------------------------------------------------------#
   #--testing for Normality--#
   SHAPIRO_NORM <- try(shapiro.test(ANALYTE_SUB$MEAN_REPORTED)$p.value, silent=TRUE)
   SHAPIRO_LOG <- try(shapiro.test(log(ANALYTE_SUB$MEAN_REPORTED))$p.value, silent=TRUE)

   #--Returns dummy value if test is an error--#
   if(class(SHAPIRO_NORM) =="try-error") SHAPIRO_NORM <- -999
   if(class(SHAPIRO_LOG) =="try-error") SHAPIRO_LOG <- -999
   #--------------------------------------------------------------------------#

   #--------------------------------------------------------------------------#
   #--Outliers Analysis--#
   GRUBBS <- grubbs.test(ANALYTE_SUB$MEAN_REPORTED, type=10, opposite=FALSE, two.sided=FALSE)
   GRUBBSLOG <- grubbs.test(log(ANALYTE_SUB$MEAN_REPORTED), type=10, opposite=FALSE, two.sided=FALSE)
   GRUBBS_OUT = GRUBBS$alt
   GRUBBSLOG_OUT = GRUBBSLOG$alt
   #--------------------------------------------------------------------------#

   #--------------------------------------------------------------------------#
   #--Summary of Analysis--#
   OUTPUT_SUM <- data.frame(Site = unique(ANALYTE_SUB$RCRA.Site),
               UpgradientWells = nrow(UP[RCRA.Site == SITES[i]]),
               Analyte = ANALYTES[j],
               Start_Date = min(STARTDATE$START_DATE),
               End_Date = max(ANALYTE_SUB$EVENT),
               N = N,
               Prob_ND = mean(ANALYTE_SUB$NDS),
               Mean = MN,
               Median = median(ANALYTE_SUB$MEAN_REPORTED),
               SD = SD,
               CV = SD/MN,
               Max = max(ANALYTE_SUB$MEAN_REPORTED),
               NCOMP = NCOMP_SUB,
               ACrit = ALPHA_CRIT,
               TCrit = TCRIT,
               CMDW = CMDW,
               CMUP = CMUP,
               Shapiro_Norm = SHAPIRO_NORM,
               Shaprio_Log = SHAPIRO_LOG,
```

```
                Grubbs = GRUBBS$p.value,
                Grubbs_log = GRUBBSLOG$p.value,
                Grubbs_Out = GRUBBS_OUT,
                Grubbslog_Out = GRUBBSLOG_OUT,
                ND_Method = ND_METHOD)
    #----------------------------------------------------------------------------#

    #----------------------------------------------------------------------------#
    #--Output Matrix--#
    OUT  <- rbind(OUT, OUTPUT_SUM)
    #----------------------------------------------------------------------------#
  }
  #----------------------------------------------------------------------------#
}
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--New Start Dates for 216-S-10 Pond and Ditch--#
SDATES2 <- readWorksheet(RCRA,sheet = 'Start_Dates_Additional',header = TRUE)
SDATES2$START_DATE <- as.Date(as.POSIXct(SDATES2$START_DATE,format='%m/%d/%Y',tz='UTC'))
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Subsetting dataset by Site--#
SUB <- DATA_SUM[RCRA.Site == '216-S-10 Pond and Ditch']
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Number of comparisons--#
NCOMP_SUB <- subset(NCOMP,Site == '216-S-10 Pond and Ditch')$NCOMP * length(ANALYTES)
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Analyte Loop--#
for(j in 1:length(ANALYTES)){

  #----------------------------------------------------------------------------#
  #--Subset Data by Analyte--#
  ANALYTE_SUB <- SUB[ANALYTE == ANALYTES[j]]
  #----------------------------------------------------------------------------#

  #----------------------------------------------------------------------------#
  #--Subset data by start date--#
  STARTDATE <- subset(SDATES2,SDATES2$RCRA.Site == '216-S-10 Pond and Ditch' & SDATES2$ANALYTE == ANALYTES[j])
  ANALYTE_SUB <- subset(ANALYTE_SUB,ANALYTE_SUB$EVENT >= min(STARTDATE$START_DATE))
  #----------------------------------------------------------------------------#

  #----------------------------------------------------------------------------#
  #--Skip Analytes with No Data--#
  if(nrow(ANALYTE_SUB) == 0){
    next
  }
  #----------------------------------------------------------------------------#

  #----------------------------------------------------------------------------#
  #--Handling Non-detects--#
  if(mean(ANALYTE_SUB$NDS) <= 0.1){

    ND_METHOD <- "HALF.RL"
    ANALYTE_SUB$ADJVAL <- ifelse(ANALYTE_SUB$NDS ==
TRUE,ANALYTE_SUB$MEAN_REPORTED/2,ANALYTE_SUB$MEAN_REPORTED)
    MN <- mean(ANALYTE_SUB$ADJVAL)
    SD <- sd(ANALYTE_SUB$ADJVAL)

  } else {

    ND_METHOD <- "Kaplan-Meier"

    KM <- enparCensored(ANALYTE_SUB$MEAN_REPORTED,
              ANALYTE_SUB$NDS)
```

```
  MN <- as.numeric(KM$parameters[1])
  SD <- as.numeric(KM$parameters[2])

}
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Determining Critical Values for Student t-Test--#
if(ANALYTES[j] == 'pH Measurement'){
  ALPHA_CRIT <- (ALPHA/NCOMP_SUB)/2
}            else {
  ALPHA_CRIT <- ALPHA/NCOMP_SUB
}

N <- nrow(ANALYTE_SUB)
TCRIT <- qt(ALPHA_CRIT,df=N-1,lower.tail=FALSE,log=FALSE)
CMUP <- MN + TCRIT *SD * sqrt(1 + 1/N)
CMDW <- MN - TCRIT *SD * sqrt(1 + 1/N)
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--testing for Normality--#
SHAPIRO_NORM <- try(shapiro.test(ANALYTE_SUB$MEAN_REPORTED)$p.value, silent=TRUE)
SHAPIRO_LOG <- try(shapiro.test(log(ANALYTE_SUB$MEAN_REPORTED))$p.value, silent=TRUE)

#--Returns dummy value if test is an error--#
if(class(SHAPIRO_NORM) =="try-error") SHAPIRO_NORM <- -999
if(class(SHAPIRO_LOG) =="try-error") SHAPIRO_LOG <- -999
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Outliers Analysis--#
GRUBBS <- grubbs.test(ANALYTE_SUB$MEAN_REPORTED, type=10, opposite=FALSE, two.sided=FALSE)
GRUBBSLOG <- grubbs.test(log(ANALYTE_SUB$MEAN_REPORTED), type=10, opposite=FALSE, two.sided=FALSE)
GRUBBS_OUT = GRUBBS$alt
GRUBBSLOG_OUT = GRUBBSLOG$alt
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Summary of Analysis--#
OUTPUT_SUM <- data.frame(Site = unique(ANALYTE_SUB$RCRA.Site),
                UpgradientWells = nrow(UP[RCRA.Site == SITES[i]]),
                Analyte = ANALYTES[j],
                Start_Date = min(STARTDATE$START_DATE),
                End_Date = max(ANALYTE_SUB$EVENT),
                N = N,
                Prob_ND = mean(ANALYTE_SUB$NDS),
                Mean = MN,
                Median = median(ANALYTE_SUB$MEAN_REPORTED),
                SD = SD,
                CV = SD/MN,
                Max = max(ANALYTE_SUB$MEAN_REPORTED),
                NCOMP = NCOMP_SUB,
                ACrit = ALPHA_CRIT,
                TCrit = TCRIT,
                CMDW = CMDW,
                CMUP = CMUP,
                Shapiro_Norm = SHAPIRO_NORM,
                Shaprio_Log = SHAPIRO_LOG,
                Grubbs = GRUBBS$p.value,
                Grubbs_log = GRUBBSLOG$p.value,
                Grubbs_Out = GRUBBS_OUT,
                Grubbslog_Out = GRUBBSLOG_OUT,
                ND_Method = ND_METHOD)
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Output Matrix--#
OUT  <- rbind(OUT, OUTPUT_SUM)
#----------------------------------------------------------------------------#
```

```
}
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Export output matrix--#
FILE <- paste0('Output/Runs - 3-7-2019/HANFORD-CRITICAL_MEANS-INTERWELL_2018_',as.character(Sys.Date()),'.csv')
write.table(OUT,FILE, sep = ",", row.names = FALSE)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--This script calculates the Critical Means for Hanford RCRA Areas for CY-2019
#--Intrawell Method for LLWMA-2 and LLWMA-3
#--Script written by: E. DiFilippo (SSP&A): modified from ECF-HANFORD-13-0013, REV. 0
#--R Version: 3.1.3
#--Package data.table version: 1.10.4-3
#--Package EnvStats version: 2.1.0
#--Package outliers version: 0.14
#--Package lawstat version: 3.2
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Set working directory--#
DIR <- 'Z:/1639_HANFORD_FY2019/66007-23 RCRA Regulatory Support GW Data Evaluation/RCRA Critical Means'
setwd(DIR)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Load Libraries--#
library(data.table)
library(EnvStats)
library(XLConnect)
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import Interwell Dataset--#
load('Data/Runs - 3-7-2019/Intrawell.RData')
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Import Staring Dates--#
RCRA <- loadWorkbook('Data/Received from Art Lee - 12-27-2018/RCRA_Wells_for_2019_CM_Calcs - Revised 3-7-2019.xlsx')
SDATES <- readWorksheet(RCRA,sheet = 'Start_Dates_Intrawell',header = TRUE)
SDATES$START_DATE <- as.Date(as.POSIXct(SDATES$START_DATE,format='%m/%d/%Y',tz='UTC'))
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Subset for Upgradient Wells--#
SITES <- as.data.table(INTRA$INTRASITES)
DATA <- INTRA$DATA
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Remove Data Based on Review Qualifiers and Collection Purpose--#
DATA <- DATA[REVIEW_RM == FALSE]
DATA <- DATA[COLLECT_RM == FALSE]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Merge Sample Data and Well Info Data--#
setkey(DATA,NAME)
setkey(SITES,Well.Name)
DATA <- SITES[DATA]
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Check units consistancy--#
UNITCHECK <- aggregate(DATA$UNIT, by=list(DATA$ANALYTE),
                  function(x){length(unique(x))})
if(max(UNITCHECK$x) > 1){
  stop("Different units - check")
```

```
}
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Calculate the number of comparison made at each site--#
NCOMP <- data.table(as.data.frame(table(INTRA$INTRASITES$RCRA.Site)))
setnames(NCOMP,colnames(NCOMP),c("Site","NCOMP"))
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Create lists of sites--#
SITES <- sort(unique(DATA$RCRA.Site))
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Calculating Summary Statistics for each Well by Sample Date
DATA_SUM <- data.table(DATA)[,list(N_NDS=sum(NDS),
                                   N_SAMPLE=length(NDS),
                                   PROPND=sum(NDS)/length(NDS),
                                   MEAN_REPORTED=mean(VAL),
                                   DETLIMIT=min(VAL),
                                   MEAN_DETECTS=mean(VAL[!NDS]),
                                   SD_DETECTS=sd(VAL[!NDS]),
                                   CV_DETECTS=sd(VAL[!NDS])/mean(VAL[!NDS])),
                             by=c("RCRA.Site","Loc","Well.Name","EVENT","ANALYTE")]
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Identify nonDetects; TRUE indicates nondetect for Summary Dataset--#
DATA_SUM$NDS <- ifelse(DATA_SUM$PROPND == 1,TRUE,FALSE)
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Calculate Summary Statistics for Each Site including NDs--#
DATA_SUM_COUNTS <- data.table(DATA_SUM)[,list(ALL_NDS=sum(PROPND==1),
                                        N_SAMPLES=length(PROPND),
                                        PROPNDALL=round(100*sum(PROPND==1)/length(PROPND))),
                                     by=c("ANALYTE","RCRA.Site")]
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Creating data frame to place results for exporting--#
OUT <- NULL
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Loop through Each Site--#
for(j in 1:length(SITES)){

  #-----------------------------------------------------------------------------#
  #--Subset by Site--#
  SUB <- DATA_SUM[RCRA.Site == SITES[j]]
  #-----------------------------------------------------------------------------#

  #-----------------------------------------------------------------------------#
  #--Create List of Analytes and Wells--#
  ANALYTES <- sort(unique(SUB$ANALYTE))
  WELLS <- sort(unique(SUB$Well.Name))
  #-----------------------------------------------------------------------------#

  #-----------------------------------------------------------------------------#
  #--Loop through Each Analyte--#
  for(i in 1:length(ANALYTES)) {

    #-----------------------------------------------------------------------------#
    #--Subsetting dataset by Site--#
    ANALYTE_SUB <- SUB[ANALYTE==ANALYTES[i]]
    if(SITES[j] == "116-N-3 Crib" & ANALYTES[i] == "Specific Conductance"){
      ANALYTE_SUB <- ANALYTE_SUB[MEAN_REPORTED < 700]
    }
```

```
    #------------------------------------------------------------------------------#

    #------------------------------------------------------------------------------#
    #--Subset data by start date--#
    STARTDATE <- subset(SDATES,SDATES$RCRA.Site == SITES[j] & SDATES$ANALYTE == ANALYTES[i])
    ANALYTE_SUB <- subset(ANALYTE_SUB,ANALYTE_SUB$EVENT >= min(STARTDATE$START_DATE))
    if(nrow(ANALYTE_SUB) == 0){
      next
    }
    #------------------------------------------------------------------------------#

    #------------------------------------------------------------------------------#
    #--Determine Methodology for Handling Non-Detects--#
    if(mean(ANALYTE_SUB$NDS) <= 0.1){

      #------------------------------------------------------------------------------#
      #--Set ND Method and Calculate 1/2 NDs--#
      ND_METHOD <- "HALF.RL"
      ANALYTE_SUB$ADJVAL <- ifelse(ANALYTE_SUB$NDS ==
TRUE,ANALYTE_SUB$MEAN_REPORTED/2,ANALYTE_SUB$MEAN_REPORTED)
      #------------------------------------------------------------------------------#

      #------------------------------------------------------------------------------#
      #--Calculate Statistics for Pooled Variance--#
      if (ANALYTES[i] == 'pH Measurement'){
        SD <- data.table(ANALYTE_SUB)[,list(N=length(NDS),
                                    SD=sd(ADJVAL),
                                    VAR=var(ADJVAL)),
                                 by=c('Well.Name')]
      } else {
        SD <- data.table(ANALYTE_SUB)[,list(N=length(NDS),
                                    SD=sd(log(ADJVAL)),
                                    VAR=var(log(ADJVAL))),
                                 by=c('Well.Name')]
      }
      #------------------------------------------------------------------------------#

      #------------------------------------------------------------------------------#
      #--Calculate Summary Statistics by Well--#
      STATS <- data.table(ANALYTE_SUB)[,list(MEAN=mean(ADJVAL),
                                     VAR=var(ADJVAL),
                                     SD=sd(ADJVAL),
                                     MEAN_Log=mean(log(ADJVAL)),
                                     VAR_Log=var(log(ADJVAL)),
                                     SD_Log=sd(log(ADJVAL)),
                                     N=length(ADJVAL),
                                     PNDS=mean(ANALYTE_SUB$NDS),
                                     MAX=max(MEAN_REPORTED,na.rm=TRUE),
                                     Median=median(MEAN_REPORTED, na.rm=TRUE),
                                     Max.date=max(EVENT, na.rm=TRUE)),
                                 by=c('RCRA.Site','Well.Name','ANALYTE')]
      #------------------------------------------------------------------------------#

    } else {

      #------------------------------------------------------------------------------#
      #--Set ND Method--#
      ND_METHOD <- "Kaplan-Meier"
      #------------------------------------------------------------------------------#

      #------------------------------------------------------------------------------#
      #--Set-up SUmmary Tables--#
      STATS <- NULL
      SD <- NULL
      #------------------------------------------------------------------------------#

      #------------------------------------------------------------------------------#
      #--Loop Through Each Well--#
      for(k in 1:length(WELLS)){
```

```
#----------------------------------------------------------------------------#
#--Subset by Well--#
WELL_SUB <- subset(ANALYTE_SUB,ANALYTE_SUB$Well.Name == WELLS[k])
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Calculate Natural Log Values--#
if (ANALYTES[i] == 'pH Measurement'){
  WELL_SUB$ADJVAL <- WELL_SUB$MEAN_REPORTED
} else {
  WELL_SUB$ADJVAL <- log(WELL_SUB$MEAN_REPORTED)
}
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Determine Statistics--#
if(mean(WELL_SUB$NDS) == 0){

  N <- length(WELL_SUB$NDS)
  MEAN <- mean(WELL_SUB$MEAN_REPORTED)
  MEAN_Log <- mean(log(WELL_SUB$MEAN_REPORTED))
  Sigma <- sd(WELL_SUB$MEAN_REPORTED)
  Sigma_Log <- sd(log(WELL_SUB$MEAN_REPORTED))
  VAR <- var(WELL_SUB$MEAN_REPORTED)
  VAR_Log <- var(log(WELL_SUB$MEAN_REPORTED))
  Median <- median(WELL_SUB$MEAN_REPORTED)
  Median_log <- median(log(WELL_SUB$MEAN_REPORTED))
  MaxDate <- max(WELL_SUB$EVENT,na.rm=TRUE)

} else if(mean(WELL_SUB$NDS) > 0 & mean(WELL_SUB$NDS) <= 0.1){

  WELL_SUB$ADJVAL <- ifelse(WELL_SUB$NDS == TRUE,WELL_SUB$MEAN_REPORTED/2,WELL_SUB$MEAN_REPORTED)

  N <- length(WELL_SUB$NDS)
  MEAN <- mean(WELL_SUB$ADJVAL)
  MEAN_Log <- mean(log(WELL_SUB$ADJVAL))
  Sigma <- sd(WELL_SUB$ADJVAL)
  Sigma_Log <- sd(log(WELL_SUB$ADJVAL))
  VAR <- var(WELL_SUB$ADJVAL)
  VAR_Log <- var(log(WELL_SUB$ADJVAL))
  Median <- median(WELL_SUB$ADJVAL)
  Median_log <- median(log(WELL_SUB$ADJVAL))
  MaxDate <- max(WELL_SUB$EVENT,na.rm=TRUE)

} else {

  KM <- enparCensored(WELL_SUB$MEAN_REPORTED,
                 WELL_SUB$NDS)
  KM_log <- enparCensored(WELL_SUB$ADJVAL,
                   WELL_SUB$NDS)

  N <- length(WELL_SUB$NDS)
  MEAN <- as.numeric(KM$parameters[1])
  MEAN_Log <- as.numeric(KM_log$parameters[1])
  Sigma <- as.numeric(KM$parameters[2])
  Sigma_Log <- as.numeric(KM_log$parameters[2])
  VAR <- Sigma^2
  VAR_Log <- Sigma_Log^2
  Median <- median(WELL_SUB$MEAN_REPORTED)
  Median_log <- median(log(WELL_SUB$MEAN_REPORTED))
  MaxDate <- max(WELL_SUB$EVENT,na.rm=TRUE)

}
#----------------------------------------------------------------------------#

#----------------------------------------------------------------------------#
#--Calculate Statistics for Pooled Variance--#
if (ANALYTES[i] == 'pH Measurement'){
  S_SUB <- data.table(Well.NAME=WELLS[k],
                 N=N,
```

```
              MEAN=MEAN,
              SD=Sigma,
              VAR=VAR)
  } else {
   S_SUB <- data.table(Well.NAME=WELLS[k],
              N=N,
              MEAN=MEAN_Log,
              SD=Sigma_Log,
              VAR=VAR_Log)
  }
  SD <- rbind(SD,S_SUB)
  #-----------------------------------------------------------------------------#

  #-----------------------------------------------------------------------------#
  #--Calculate Summary Statistics--#
  STATS_SUB <- data.table(RCRA.Site = SITES[j],
               Well.Name = WELLS[k],
               ANALYTE = ANALYTES[i],
               MEAN = MEAN,
               VAR = VAR,
               SD = Sigma,
               MEAN_Log = MEAN_Log,
               VAR_Log = VAR_Log,
               SD_Log = Sigma_Log,
               N = length(WELL_SUB$MEAN_REPORTED),
               PNDS=mean(WELL_SUB$NDS),
               MAX=max(WELL_SUB$MEAN_REPORTED,na.rm=TRUE),
               Median=Median,
               Max.date=MaxDate)
  STATS <- rbind(STATS,STATS_SUB)
  #-----------------------------------------------------------------------------#

 }
 #-----------------------------------------------------------------------------#

}
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Pooled Variance--#
if(SITES[j] == 'LLWMA-2' & ANALYTES[i] %in% c('pH Measurement') |
  SITES[j] == 'LLWMA-3' & ANALYTES[i] %in% c('pH Measurement','Specific Conductance')){

 SD$VAR2 <- (SD$N-1)*SD$VAR
 VAR_POOLED <- sum(SD$VAR2,na.rm=TRUE)/(nrow(ANALYTE_SUB)-length(unique(ANALYTE_SUB$Well.Name)))
 SD_POOLED <- sqrt(VAR_POOLED)
 N_POOLED <- length(ANALYTE_SUB$MEAN_REPORTED)
 DF_POOLED <- N_POOLED-length(unique(ANALYTE_SUB$Well.Name))

}
#-----------------------------------------------------------------------------#

#-----------------------------------------------------------------------------#
#--Summary Statistics--#
if(SITES[j] == 'LLWMA-2' & ANALYTES[i] %in% c('pH Measurement') |
  SITES[j] == 'LLWMA-3' & ANALYTES[i] %in% c('pH Measurement','Specific Conductance')){

 STATS$POOLED_SD <- SD_POOLED
 STATS$POOLED_N <- N_POOLED
 STATS$ALPHA <- 0.01
 STATS$DF_POOLED <- DF_POOLED
 if(ANALYTES[i] == 'pH Measurement'){
  STATS$TCRIT <- qt(1-STATS$ALPHA/2,df=DF_POOLED)
 } else {
  STATS$TCRIT <- qt(1-STATS$ALPHA,df=DF_POOLED)
 }


} else {
```

```
      STATS$POOLED_SD <- NA
      STATS$POOLED_N <- NA
      STATS$ALPHA <- 0.01
      STATS$DF_POOLED <- NA

    }
    #------------------------------------------------------------------------------#

    #------------------------------------------------------------------------------#
    #--Calculate Critical Means--#
    if(SITES[j] == 'LLWMA-2' & ANALYTES[i] %in% c('pH Measurement') |
       SITES[j] == 'LLWMA-3' & ANALYTES[i] %in% c('pH Measurement','Specific Conductance')){

      if (ANALYTES[i] == 'pH Measurement'){

        STATS$CMUP_LOG <- NA
        STATS$CMDW_LOG <- NA
        STATS$CMUP <- STATS$MEAN + STATS$TCRIT * STATS$POOLED_SD * sqrt(1 + 1/STATS$POOLED_N)
        STATS$CMDW <- STATS$MEAN - STATS$TCRIT * STATS$POOLED_SD * sqrt(1 + 1/STATS$POOLED_N)

      } else {

        STATS$CMUP_LOG <- STATS$MEAN_Log + STATS$TCRIT * STATS$POOLED_SD * sqrt(1 + 1/STATS$POOLED_N)
        STATS$CMDW_LOG <- NA
        STATS$CMUP <- exp(STATS$CMUP_LOG)
        STATS$CMDW <- exp(STATS$CMDW_LOG)

      }

    } else {

      if(ANALYTES[i] == 'pH Measurement'){
        ALPHA_CRIT <- 0.01/2
      }          else {
        ALPHA_CRIT <- 0.01
      }

      STATS$TCRIT <- qt(ALPHA_CRIT,df=STATS$N-1,lower.tail=FALSE,log=FALSE)
      STATS$CMUP_LOG <- NA
      STATS$CMDW_LOG <- NA
      STATS$CMUP <- STATS$MEAN + STATS$TCRIT *STATS$SD * sqrt(1 + 1/STATS$N)
      STATS$CMDW <- STATS$MEAN - STATS$TCRIT *STATS$SD * sqrt(1 + 1/STATS$N)

    }
    OUT <- rbind(OUT,STATS)
    #------------------------------------------------------------------------------#

  }
  #------------------------------------------------------------------------------#

}
#------------------------------------------------------------------------------#

#------------------------------------------------------------------------------#
#--Export output matrix--#
FILE <- paste0('Output/Runs - 3-7-2019/HANFORD-CRITICAL_MEANS-INTRAWELL_2018_',as.character(Sys.Date()),'.csv')
write.table(OUT,FILE, sep = ",", row.names = FALSE)
save(OUT,file='Output/Runs - 3-7-2019/IntrawellTest_CMs.Rdata')
save(DATA_SUM,file='Output/Runs - 3-7-2019/IntrawellTest_SummedData.RData')
#------------------------------------------------------------------------------#
```